DOCUMENT RESUME

ED 175 550 PS 010 830

AUTHOR FitzGibbon, Ann
TITLE The Self-Concept Target Game: A Comprehensive Report.
INSTITUTION Far West Lab. for Educational Research and Development, San Francisco, Calif.
PUB DATE Aug 73
NOTE 72p.

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Aspiration; Blacks; Caucasians; Elementary Education; *Elementary School Students; Grade 2; Grade 3; *Psychological Tests; *Racial Differences; Research; *Self Concept; *Sex Differences; Test Reliability; Test Validity
IDENTIFIERS *Responsive Environment Model; Self Concept Target Game

ABSTRACT

A Self Concept Target Game (SCTG) which was designed to measure the level of aspiration of second or third grade children who had been exposed to at least two years of the Responsive Model Follow Through (RMFT) Program, and the field testing of the game are described in this report. The game is played by throwing and pushing a bean bag along a plastic board to hit targets of varying difficulty, after indicating which target area one expects to hit. Three field tests of the game were conducted. Results of the first test indicated that the game had validity since the results correlated with teachers' choices of children high and low in self concept. The second test indicated that RMFT children performed better than children not exposed to the special program and that the children with the most realistic expectations were those taught by black teachers who scored highest on a Responsive Observation Schedule. The third and largest field test included 553 second and third graders in five school districts throughout the country. Total group and school district results showed that second graders did better than third graders, that blacks were less realistic on their expectations than whites, and that boys were less realistic than girls. RMFT program enrollment did not seem to make a difference in score. Descriptive total and individual item scores and correlations among the items and with the total score are given separately for each school district and for the total group. Conditions for use of the SCTG by researchers are specified. (Author/RH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED175550

PS

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

THE SELF-CONCEPT TARGET GAME

A Comprehensive Report

August 31, 1973.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Far West Lab. for Educational Research and Development
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

PS010830

TABLE OF CONTENTS

	Page
Description of the Game	1
The Target Population	3
Instruction for Administering	4
Giving Feedback to the Subject	6
First Field Test	8
Second Field Test	8
Third Field Test	9
1. Distributional Characteristics of TGS	10
2. Other Valuables Associated with TGS	12
3. Characteristics of T. G. Items	19
4. Reliability of the T.G	19
5. Validity of TGS	24
6. Problems in Interpreting Score	24
a) Group Testing	24
b) Physical Format	25
7. Summary	28
Appendix A	32
Appendix B	33

THE SELF CONCEPT TARGET GAME

A Comprehensive Report
by Ann FitzGibbon

Description of the Game

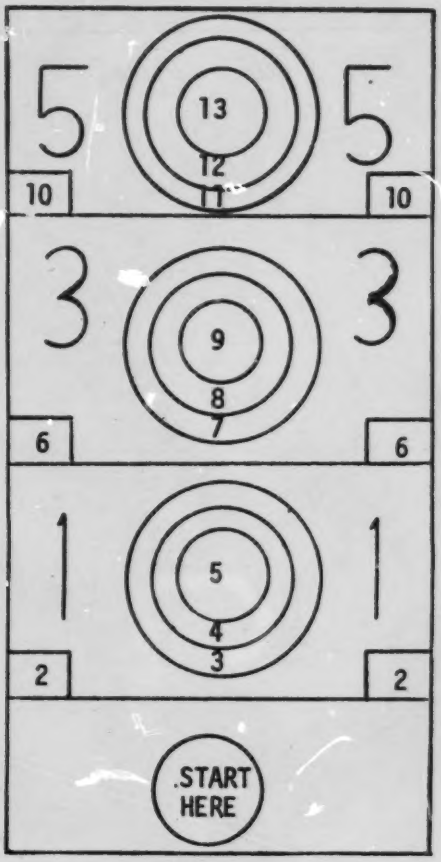
From time to time there has appeared in the literature an empirical relationship between self esteem or a healthy self concept and level of aspiration (Mitchell, 1959; Borislow, 1962; Harrison, 1971;and Mukherjee, 1965). Certainly it is to be expected that one of the functions of a healthy concept of self is the setting of goals which do not exceed the boundaries of reality. Therefore, when the healthy self concept was defined for the Responsive Model child, it included:

> "...that such a child be able to make good estimates of his ability to perform a given task, and be willing to take reasonable risks of failure... (Nimnicht, 1970)"

This definition refelcts the philosophy of the Responsive Model Early Education Program which tries to enstill in children the "Hope of Success" Motivation and to help them eliminate "Fear of Failure."

Since the concept of self is made up of many areas, a global approach at measurement has been abandoned. Instead, several specific tasks are being designed, each attempting to measure some area of the Responsive Model child's self concept. The Target Game is one of these tasks.

The game consists of a plastic strip 78" X 20" with three main target areas:



The game is played by throwing or pushing a bean bag along the plastic board, attempting to hit a target with a particular value. The large box nearest the starting circle has a value of one, the middle box a value of 3, and the box farthest from the subject a value of 5. Within each large area are smaller targets, each with more points to earn. Since the smaller targets are harder to hit, they are worth more points. In order to earn points a child must say "I'm going to hit ___" or "I'm going to try for ___" and then the bean bag must land somewhere within that area. If the subject misses, he earns no points on that trial.

While the object of the game as explained to the subjects is to win as many points or as high a score as possible, we are primarily interested in how children revised their choices based on feedback from successive trials at the game. For this reason, each subject's score is determined by subtracting the value of his try (guess) on any one trial from his actual performance (hit) on the previous trial. These values are summed (ignoring the sign) and the result is the Self Concept Discrepancy score. Since smaller discrepancies between estimates and past performances result in lower scores, it follows that a lower (rather than a higher) score is more desirable.

This discrepancy score or "goal discrepancy" score is a now standard quantitative technique developed by Frank (1935) for the experimental measurement of level of aspiration and described by Redfearn (1969). Past performance has been demonstrated by several investigators to be a valid predictor of level of aspiration. Frank (1935) found that the difference between goals and median level of past performance to be highly reliable and independent of the task in which it is determined. Sears (1940) found that self confident, successful 4th, 5th, and 6th grade children determined their goals by past achievement; and Birney, Burdick and Teevan (1969) showed that subjects with High Hope of Success motivation made aspiration estimates which reflected their actual performance while High Fear of Failure subjects gave estimates out of touch with their performance.

The Target Population

The Self Concept Target Game was designed to measure second or third grade children who have been exposed to at least two years of

the Responsive Model Follow Through program. The game is not suitable for younger children, and it is felt that at least two years exposure to the special program is necessary for a perceptible change in the level of aspiration to occur. Since the population of the Responsive Model program is approximately 60% Black, 30% Chicano, and 10% lower-class white, a change in the level of aspiration of minority group children is a realistic goal. Goldblatt and Tyson (1962) found that 4th, 5th and 6th grade white children had higher levels of aspiration than Black or Puerto Rican children. Guggenheim (1967) reported that Black children's level of aspiration drops after experience of failure although the level of aspiration of Blacks and Whites was similar initially. Trowbridge states (1972) that low SES children from 9 to 14 years have lower levels of aspiration, expect less of themselves and perceive parents, teachers and peers as expecting less of them than do Middle SES children. Brim et.al. (1969) found that White children had more realistic estimates of their abilities than Black children, while Harrison found (1971) that performance - discrepant goal setting is more charactistic of children with a long history of failure than of those without one.

Instructions for Administering the Self Concept Target Game

The Target Game may be administered to groups of five children if necessary; individual administration is desirable if time permits. (The additional instructions for groups administration are in parentheses, otherwise instructions are the same for both individuals and groups of five).

Spread out plastic strip so that it is entirely visible to subject(s), and say (YOU HAVE A CHOICE. YOU CAN EITHER STAY HERE AND PLAY THE GAME, OR YOU CAN GO BACK TO THE CLASSROOM. IF YOU DECIDE TO STAY AND PLAY THE GAME THIS IS WHAT WE'LL DO. I WILL EXPLAIN HOW TO PLAY THE GAME. ONE OF YOU WILL TAKE HIS TURN, WHILE THE REST OF YOU SIT QUIETLY AT THAT TABLE UNTIL IT IS YOUR TURN. YOU MUST SIT QUIETLY SO YOU DON'T DISTURB THE ONE WHO IS PLAYING. DOES ANYONE WANT TO GO BACK TO THE CLASSROOM?) THIS IS A GAME WHERE WE TRY TO EARN AS MANY POINTS AS WE CAN BY EITHER PUSHING (DEMONSTRATE) OR THROWING (DEMONSTRATE) THE BEAN BAG ALONG THE PLASTIC BOARD. THIS (POINT) IS THE ONE BOX, AND IF IT LANDS ANYWHERE IN HERE, I EARN ONE POINT. THIS (POINT) IS THE THREE BOX. IF I SAY MY BEAN BAG IS GOING TO LAND IN THE THREE BOX AND IT LANDS ANYWHERE IN THERE, I EARN THREE POINTS. WHAT IS THIS (POINT)? THAT'S RIGHT, THE FIVE BOX. IF I TRY TO HIT ANYWHERE IN THE FIVE BOX AND I DO HIT IT, I EARN FIVE POINTS. REMEMBER THAT IF I MISS WHAT I SAY I AM GOING TO HIT, THEN I DON'T EARN ANY POINTS AT ALL. YOU MUST HIT WHAT YOU SAY YOU ARE GOING TO HIT OR YOU DON'T EARN ANY POINTS. THESE SMALLER TARGETS (POINT TO NUMBERS) ARE WORTH MORE POINTS, BUT THEY ARE HARDER TO HIT. IF I SAY I AM GOING TO HIT THE NINE CIRCLE (DEMONSTRATE) AND IT LANDS IN THERE, I EARN NINE POINTS, BUT IF I MISS, THEN I DON'T EARN ANY POINTS AT ALL. YOU WANT TO EARN AS MANY POINTS AS YOU CAN, SO CHOOSE YOUR TARGETS CAREFULLY. THE SMALLER ONES ARE WORTH MORE POINTS, BUT THEY ARE HARDER TO HIT. (DOES ANYONE WANT TO GO BACK TO THE CLASSROOM? WHO WANTS TO GO FIRST?")

As child approaches the game, say,

WHAT WOULD YOU LIKE TO TRY FOR FIRST? (Mark score sheet in accordance with child's choice.) After child throws or pushes the bean bag, say, YOU WANTED TO HIT THE ____, AND YOU HIT THE ____. (Mark score sheet).

Instructions for Administering (cont'd)

If child hit target he chose, say, YOU EARNED ____ POINTS. WHAT WOULD YOU LIKE TO TRY FOR NEXT?

If child hit another target than the one he chose, say, YOU TRIED FOR THE ____ AND YOU HIT THE ____. YOU DIDN'T EARN ANY POINTS THIS TIME, BUT WHAT WOULD YOU LIKE TO TRY FOR NEXT? (Occasionally encourage child when he misses target by saying YOU WERE VERY CLOSE.)

Continue asking each child to choose his targets until all 11 trials are exhausted. Circle each choice child makes on each trial, and each target he hits. Fill in information requested at top of sheet, and if possible, any verbal behavior at the bottom.

Giving Feedback to the Subject

Children often asked during the field tests how they were doing on the game, or how many points they had earned. The examiner was instructed to make comments such as "Good, you hit the 3 box and earned 3 points," or "You earned 6 points for hitting the six box." "You are all doing very well." If a child asks about a numerical score, he was given a value in agreement with the one described in Instructions for Administering, and not the level of Aspiration Discrepancy Score. If the examiner wished, she could record in the left margin of the Score Sheet each time a chosen target is hit, and the value obtained, so that she could quickly sum these if the child asked.

To inquiries of "Who got the most points?" the examiner replied "I have not added all the scores which I will do later. You all did very well." Level of Aspiration Discrepancy Scores were calculated when all the children had been tested.

7

SELF CONCEPT TARGET GAME SCORE SHEET

Name_____ School_____ Date _____

Teacher_____ Grade____ City_____ B. D._____

FT_____ Other _____ Race_____ Sex _____ Age _____

Trial		Circle Number in Each Row	
	Chooses	1 2 3 3 4 5 5 6 7 8 9 10 11 12 13	
	Hits	1 2 3 3 4 5 5 6 7 8 9 10 11 12 13	
	Chooses	1 2 3 3 4 5 5 6 7 8 9 10 11 12 13	(Hits - Chooses)___
1	Hits	1 2 3 3 4 5 5 6 7 8 9 10 11 12 13	
	Chooses	1 2 3 3 4 5 5 6 7 8 9 10 11 12 13	(Hits - Chooses)___
2	Hits	1 2 3 3 4 5 5 6 7 8 9 10 11 12 13	
	Chooses	1 2 3 3 4 5 5 6 7 8 9 10 11 12 13	(Hits - Chooses ___
3	Hits	1 2 3 3 4 5 5 6 7 8 9 10 11 12 13	
	Chooses	1 2 3 3 4 5 5 6 7 8 9 10 11 12 13	(Hits - Chooses ___
4	Hits	1 2 3 3 4 5 5 6 7 8 9 10 11 12 13	
	Chooses	1 2 3 3 4 5 5 6 7 8 9 10 11 12 13	(Hits - Chooses ___
5	Hits	1 2 3 3 4 5 5 6 7 8 9 10 11 12 13	
	Chooses	1 2 3 3 4 5 5 6 7 8 9 10 11 12 13	(Hits - Chooses ___
6	Hits	1 2 3 3 4 5 5 6 7 8 9 10 11 12 13	
	Chooses	1 2 3 3 4 5 5 6 7 8 9 10 11 12 13	(Hits - Chooses) ___
7	Hits	1 2 3 3 4 5 5 6 7 8 9 10 11 12 13	
	Chooses	1 2 3 3 4 5 5 6 7 8 9 10 11 12 13	(Hits - Chooses ___
8	Hits	1 2 3 3 4 5 5 6 7 8 9 10 11 12 13	
	Chooses	1 2 3 3 4 5 5 6 7 8 9 10 11 12 13	(Hits - Chooses) ___
9	Hits	1 2 3 3 4 5 5 6 7 8 9 10 11 12 13	
	Chooses	1 2 3 3 4 5 5 6 7 8 9 10 11 12 13	(Hits - Chooses) ___
10	Hits	1 2 3 3 4 5 5 6 7 8 9 10 11 12 13	
			(Sum ignoring sign) Total

Comments:

Please record child's verbal behavior during trials_____

_____

_____

_____

First Field Test

The Self Concept Target game has undergone three field tests with second and third graders. On the first and preliminary test approximately 45 children (2 classrooms) were administered a shortened form of the test (6 trials). The scoring consisted of sorting the test protocols into either real or unreal aspiration according to whether the child revised his estimates up after success and down after failure. Teachers were asked to identify the 5 children each felt to be highest in self esteem and the 5 whom each considered lowest. In one classroom, our sorting procedure identified all 5 of the teacher's high choices, and 3 of her 5 low choices. In the second classroom the procedure identified 2 of the teacher's 5 high choices, and 4 of the 5 low choices.

Second Field Test

A shortened form of the test (6 trials) was used in the second field test; this was divided between two sites: 144 black and white children in 7 integrated classrooms were tested in one site and 40 black children in 2 segregated classrooms in another site. In the first case the examiner was white, and in the second the examiner was black. Five of the 7 classrooms in Site 1 were Responsive Model Follow Through classrooms.

Mean Level of Aspiration discrepancy scores appear in Table 1. Keeping in mind that the lower the score, the more realistic the Level of Aspiration, the Follow Through children in all 5 classes out performed those in the comparison classrooms in both sites. Significant differences in favor of the Follow Through children were found in all cases except one (classroom 1). Further, children in integrated classrooms (site 1) were more realistic in their expectations on this task than were children in segregated classrooms (site 2). It is also of interest to note that the children in classrooms 2 and 5, where the scores were the most realistic, were being taught by

TABLE I

Self Concept Target Game Mean Scores by Classroom in 2 sites (6 trial administration). Second Field Test.

Site 1		N	$\bar{X}$
Follow Through			
Classroom	1	19	17.84
	2	21	12.90
	3	19	14.47
	4	18	14.50
	5	17	12.35
Non Follow Through			
Classroom	6	24	18.38
	7	26	19.08
Site 2			
Non Follow Through			
Classroom	1	19	20.53
	2	21	18.52

TABLE 2

Self Concept Target Game Mean scores by Race and Sex in 2 sites (6 trial administration). Second Field Test.

Site 1 (Integrated)

	Boys	Girls	
Black	16.94	15.41	16.10
White	16.50	14.90	15.75
Total	16.72	15.15	15.92

Site 2 (Segregated)

	Boys	Girls	
Black	19.91	18.94	19.47
White	-	-	

Black Follow Through teachers. The teacher in classroom 2 has consistently scored high on an observation instrument for effectiveness of Follow Through teachers.

The data was examined for between race and sex comparisons in the two sites; results appear in Table 2. Because we felt the Black children in the 2 sites were sufficiently far apart in test scores, a mean score for all Blacks was not computed.

There was a small but not significant difference in favor of White children in site 1 (in agreement with findings reported by Lasky, 1970); however, the difference between both Whites and Blacks in site 1 and Blacks in site 2 was significant ($p < .01$). Black children in the integrated site 1 were more realistic in their expectations than were Black children in the segregated classrooms.

In both sites girls out performed boys; in site 1 the difference in favor of girls' scores was significant ($p < .05$). While results of previous studies do not always favor girls, Lefecourt et. al (1968) found the level of aspiration of girls to be more realistic than that of boys.

Thrity-three children in two classrooms were retested with the Target Game approximately two weeks after the first testing; a test-retest coefficient of reliability of .87 was obtained, indicating the game was highly reliable.

Third Field Test

The larger part of this report is concerned with the operating characteristics of the target game discrepancy score, and the results obtained for 553 children in second and third grades tested during the spring and summer of 1972 in five school districts throughout the county. In one of the five districts, Berkeley, Calif., the children tested were in Responsive Model Follow Through program; in Omaha, Nebraska, administrators

of one ungraded catholic school allowed us to test their (approximately) second and third grade children in two classrooms to serve as a control for Follow Through subjects. In the other 3 districts, Owensboro, Goldsboro and Lebanon both Follow Through and non-Follow Through children participants in the study. The records of a small number of the total children tested are not used in all the analyses presented here (missing data, or the like), and the records of the tiny number of Mexican American, Indian, and Oriental children available from these studies are not used. Thus, the sizes of various groups discussed occasionally differ from analysis to analysis and sometimes within what looks like the same analysis.

1. Distributional Characteristics of the Target Game Scores.[1]

Means and standard deviations for the 553 children are shown in Table 3, both by district and for the total group. (Since the number of trials in the game had been increased from 6 to 11, these scores are not directly comparable to the data shown in Tables 1 and 2.) There is a significant difference between each district and every other district in mean score, except between Lebanon and Owensboro. Since a lower (less discrepant) score is more desirable than a higher one, the rank ordering of the districts is (in terms of children's performance):

Berkeley
Lebanon
Owensboro
Omaha
Goldsboro

We were disappointed to find that the control subjects (Omaha) performed significantly better than the Follow Through subjects in Goldsboro. Inspection of

[1]Appreciation is expressed to David Peizer for his analyses of the data and for help in interpreting scores.

TABLE 3

Means and Standard Deviations of Self Concept Target Game Scores (11 trials) by School District and Grade.

District	Berkeley	Goldsboro	Lebanon	Owensboro	Omaha	Total Group
n	50	129	65	260	41	553
Grade II	42	66	46	111	15	
Grade III	16	63	19	149	26	
mean	30.10	36.03	32.00	32.09	34.90	33.06
Grade II	29.26	36.24	31.65	29.77	37.87	
Grade III	34.50	35.81	32.84	33.82	33.19	
s.d.	11.53	12.27	10.98	11.18	10.89	11.56
Grade II	12.48	13.31	10.58	10.13	11.82	
Grade III	8.08	11.19	12.16	11.64	10.16	

Table 3 indicates that in Berkeley, Lebanon, and Owensboro second graders outperformed third graders (not significantly so in Lebanon) while in Omaha and Goldsboro the third graders did better.

Figure 1 is presented to help teachers and other interested school personnel in interpreting individual Target Games scores of their children. It shows the cumulative distribution function of Target Game scores based upon 545 children. Scores may be read on the horizontal axis at the top or bottom of the graph. Percent of children scoring more is read on the left vertical axis, while percent of children scoring less is read on the right vertical axis. The vertical axes are stretched out at the ends in such a way that a cumulative normal distribution function would end plotted as a straight line. The solid graph expresses population characteristics.

In order to interpret a child's score, a teacher might ask:

To what population percentile does an observed target game score of 20 correspond?

1. Find 20 on the horizontal Target Game Score axis at the top of the graph.
2. Read down to the solid graph curve.
3. Read to the left to find that 90% of children scored more; read to the right to find that 10% of children scored less.

or What Target Game Score did 40% of children score more than?

1. Find 40 on the vertical PERCENT OF CHILDREN SCORING MORE scale at the left of the graph.
2. Read across to the solid graph curve.
3. Read up or down to find that the corresponding target game score is 35.

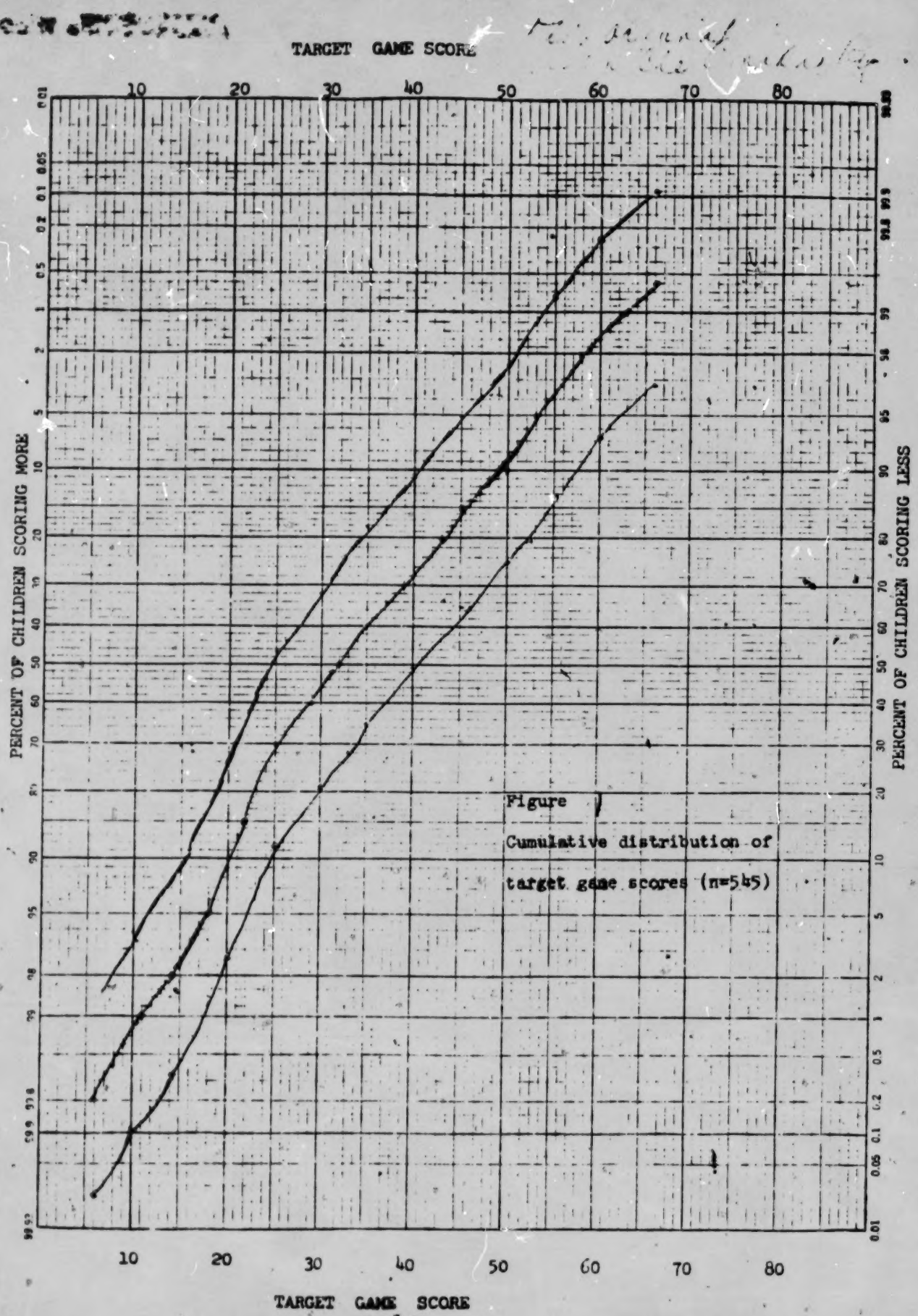

TARGET GAME SCORE
10 20 30 40 50 60 70 80
PERCENT OF CHILDREN SCORING MORE
PERCENT OF CHILDREN SCORING LESS
TARGET GAME SCORE

Figure 1

Cumulative distribution of target game scores (n=545)

The dotted graph curves relate underlying (also called 'true') percentiles to observed scores for the single child. The proportion of children bracketed by these dotted graph curves may be guessed to be 68%. For example, to what underlying percentiles may a Target game score of 50 be reasonably expected to correspond?

1. Find 50 on the horizontal Target Game score axis at the top of the graph.
2. Read down to each of the two dotted graph curves.
3. Read to the left finding more than percentiles of 2.5% and 25%.
4. It is guessed that sixty-eight percent of children having Target Game Scores of 50 have underlying more than percentiles between 2.5% and 25%, while thirty-two percent of the children have underlying more than percentiles outside these limits.

or To what Target Game Scores may an underlying percentile of 5% or less be reasonable expected to correspond?

1. Find 5 on the vertical PERCENT OF CHILDREN SCORING LESS scale at the right of the graph.
2. Read left to each of the two dotted graph curves.
3. Read up or down finding target game scores of 12 and 22.
4. It is guessed that sixty-eight percent of children having an underlying less than percentile of 5% have Target Game Scores between 12 and 22, while thirty-two percent of such children have scores outside these limits.

2. Other Variables Associated with the Target Game Score.

The preliminary analysis established that age was unrelated to score; one of the districts (Lebanon) did not report childrens' scores by race, so the

data from Lebanon was omitted in the analyses in which race was a variable. Table 4 shows how differences in race, sex, grade, and enrollment in a Follow Through program are associated with performance on the Target Game in 13 schools in the remaining 4 districts. As a whole third graders are more discrepant in choosing their targets than second graders, Blacks more than whites, boys more than girls, while being in a Follow Through program does not seem to make a difference.

Table 5 gives the same kind of information, this time broken down by district. The differences which appear for the total group in Table 4 appear in each of the districts.

Means differences in Target Game scores for grade, type of program, race and sex, each difference adjusted for the effects of the other three variables, but not for school, are shown in Table 6. Grade and sex differences continue to hold; where race differences and programmatic differences do not hold.

Parallel analyses for each of the 4 districts are shown in Table 7. Inspection of the table indicates that the districts are highly variable among themselves when means are adjusted for the effect of 3 of the variables, i.e., in Berkeley there is virtually no differences between Blacks and whites, while there is a substantial one in Goldsboro. In Berkeley, there is a substantial difference between 2nd and 3rd grades while in Goldsboro there is virtually none. However, since the standard errors are high, no conclusions can be drawn, although in the case of sex, girls consistently scored lower (more desirable) than boys.

Demographic composition and Target Game Score averages for individual schools are shown in Table 8. Reading across the table for example, it can be seen that Cragmont school in Berkeley consisted of 22 children, 4 of whom were in third grade (18%), all of whom were in Follow Through, 4 of whom were white

TABLE 4

Mean Differences of Target Game Scores for 4 Demographic Variables, Standard Errors and t' values. Each Difference Adjusted for the Effects of the Three Other Demographic Variables as well as School. (N=488)

	$\bar{X}_1$	$\bar{X}_2$	Mean[1] Diff.	S.E.	t-value
Grade (3rd-2nd)	35.08	31.17	3.91	1.20	3.26[3]
Type of Program (FT-NFT)	33.51	32.63	.88	1.38	.64
Race (White-Black)[2]	32.02	34.54	-2.52	1.18	-2.14[3]
Sex (Female-Male)	31.10	35.35	-4.25	1.00	-4.26

[1]Represents regression coefficient.

[2]Race not reported in Lebanon data.

[3]Significant $<.05$ level (2 tailed)

TABLE 5

Mean Differences of Target Game Scores for Available Demographic Variables, Adjusted for the Effects of the Other Available Demographic Variables and for School, Computed Separately for Each District.

	Berkeley	(N)	Goldsboro	(N)	Owensboro	(N)	Lebanon	(N)	Omaha	(N)
Grade: III	37.59	(16)	39.27	(63)	34.92	(148)	(1)	(19)	33.80	(26)
II	27.25	(42)	32.94	(66)	30.65	(112)		(46)	38.60	(15)
Diff.	10.34		6.33		4.27				- 4.80	
S.E.	5.67		3.44		2.31				3.67	
Type of Program:										
FT	30.10	(58)	36.62	(83)	33.40	(177)	32.35	(37)	____	(0)
NFT	____	(0)	34.96	(46)	32.44	(83)	31.54	(28)	34.90	(41)
Diff.	____		1.66		.96		.81		____	
S.E.	____		2.31		1.82		2.83		____	
Race: White	27.92	(15)	34.17	(61)	32.06	(175)	(2)		39.90	(8)
Black	30.86	(43)	37.69	(68)	35.22	(85)			33.70	(33)
Diff.	- 2.94		- 3.52		- 3.16				6.20	
S.E.	3.80		2.03		1.74				4.73	
Sex: Female	27.31	(29)	31.54	(57)	31.78	(142)	30.14	(32)	36.00	(18)
Male	32.89	(29)	39.58	(72)	34.67	(118)	33.80	(33)	34.10	(23)
Diff.	- 5.58		- 8.04		- 2.89		- 3.66		1.90	
S.E.	3.21		2.07		1.31		2.92		3.74	

(1) No school with children in both grades in Lebanon.
(2) Race not reported in Lebanon data.

TABLE 6

Mean Differences of Target Game Scores and 4 Demographic Variables, Standard Errors and t' Values. Each Mean Difference is Adjusted for the Effects of the Other Three Variables, but not for School. (N=488)

	$\bar{X}_1$	$\bar{X}_2$	Mean[1] Diff.	S.E.	t-Value
Grade (3rd-2nd)	34.26	32.06	2.20	1.09	2.03[3]
Type of Program (FT-NFT)	32.95	33.67	- .72	1.16	- .62
Race (White-Black)[2]	33.15	33.26	- .11	1.05	- .11
Sex (Female-Male)	31.06	35.38	-4.32	1.04	-4.15

[1]Represents regression coefficient.

[2]Race not reported in Lebanon Data.

[3]Significant < .05 level (2 tailed)

TABLE 7

Mean Differences of Target Game Scores for Available Demographic Variables, Adjusted for the Effects of the Other Available Demographic Variables but not School, Computed Separately for Each District.

		Berkeley (N)	Goldsboro (N)	Owensboro (N)	Lebanon (N)	Omaha (N)**
Grade:	III	36.44 (16)	36.52 (63)	33.71 (148)	*	
	II	27.69 (42)	35.56 (66)	29.95 (112		
	Diff.	8.75	.96	3.76		
	S.E.	5.94	2.65	2.33		
Type of Program:						
	FT	***	37.07 (83)	31.82 (177)	32.51 (37)	
	NFT		34.16 (46)	32.67 (83)	31.33 (28)	
	Diff.		2.91	- .85	1.18	
	S.E.		2.36	1.64	2.85	
Race:	White	29.80 (15)	34.06 (61)	33.67 (175)	****	
	Black	30.20 (43)	37.79 (68)	31.89 (85)		
	Diff.	- .40	- 3.73	1.78		
	S.E.	3.85	2.08	1.50		
Sex:	Female	26.93 (29)	31.82 (57)	30.70 (142)	29.97 (32)	
	Male	33.27 (29)	39.36 (72)	33.78 (118)	33.96 (33)	
	Diff.	- 6.33	- 7.54	- 3.08	- 3.99	
	S.E.	3.37	2.12	1.37	2.76	

*Only one grade in Lebanon.
**Only one school in Omaha.
***Only Follow Through in Berkeley.
****Race not reported in Lebanon.

TABLE 8

School, Demographic Information and Target Game Score by District (N=553).

District (School)	N	Third Grade	Follow Through	Race (White)	Sex (Female)	Mean Target Game Score	Adj. Target Game Score	S.E.
Berkeley	22	4 (18%)	22 (100%)	4 (18%)	13 (59%)	26.23	26.74	2.39
Cragmont	32	11 (34%)	32 (100%)	11 (34%)	14 (44%)	34.16	33.79	1.97
John Muir	4	1 (25%)	4 (100%)	0 (0%)	2 (50%)	27.75	27.15	5.46
Thousand Oaks	58	16 (28%)	58 (100%)	15 (26%)	29 (50%)		30.66	1.46
Goldsboro								
Walnut Street	22	0 (0%)	18 (82%)	9 (41%)	11 (50%)	35.00	36.57	2.39
Edgewood	24	0 (0%)	21 (88%)	13 (54%)	15 (63%)	39.92	42.30	2.29
Wm Street	59	39 (66%)	42 (71%)	27 (46%)	23 (39%)	36.88	35.62	1.43
East End	24	24 (100%)	12 (50%)	12 (50%)	8 (33%)	31.00	28.46	2.28
	129	63 (63%)	93 (64%)	61 (47%)	57 (44%)		35.69	.98
Omaha								
(Sacred Heart)	41	26 (63%)	0 (0%)	8 (20%)	18 (44%)	34.90	33.92	1.96
Owensboro								
Lincoln	64	44 (69%)	41 (64%)	58 (91%)	38 (59%)	34.02	34.71	1.94
Emerson	17	0 (0%)	0 (0%)	16 (94%)	11 (65%)	34.94	39.20	2.94
Franklin	74	42 (56%)	53 (71%)	36 (48%)	45 (60%)	26.81	26.90	1.26
Goodloe	40	20 (50%)	40 (100%)	7 (18%)	19 (48%)	29.20	27.96	1.81
Lee	65	43 (66%)	44 (68%)	59 (91%)	30 (46%)	36.82	37.02	1.43
	260	149 (57%)	178 (68%)	176 (67%)	143 (55%)		32.42	.61
Lebanon								
Hanover Street	19	19 (100%)	13 (68%)	*	11 (58%)	32.84	**	
School Street	24	0 (0%)	13 (54%)	*	9 (38%)	34.54	**	
Sacred Heart	22	19 (0%)	11 (50%)	*	12 (55%)	28.50	**	
	65	19 (29%)	37 (64%)	*	32 (49%)			
Total	553	273 (49%)	356 (64%)	*	279 (50%)			

*Race information not available.
**Adjustment for race and grade not possible.

(18%), 13 of whom were girls (59%), and whose average Target Game Score was 26.23. Average scores adjusted for grade, Follow Through, race, and sex, and their standard errors are also shown in Table 8.

Inspection of Table 8 suggests that schools in the same district differ from each other by wider amounts than districts were seen to differ from each other in Table 3. Significance tests confirm visual appearances: districts means are different, and school means are different within district. The standard deviation among school means within district is estimated at about 3.9. This supports appearances that schools within the same district probably differ from each other more widely than do districts differ from each other.

As was discussed in connection with the previous tables, Blacks appear to be more discrepant than Whites when scores are adjusted for school, but not otherwise; the implication is clearly that a within school effect and a between school effect are canceling each other out. The between school effect can be seen clearly in Table 8 for Owensboro. Here, Lincoln, Emerson and Lee schools have samples whose composition is 90% White or more, and target game scores are more discrepant for these schools than for the other two schools in the same district whose samples are 48% White and 18% White respectively. The group testing procedures used in obtaining target game score data confound some of the possible sources of between school variation: a hypothesis which states that a child's target game score is influenced by the racial composition of his classroom or school and a hypothesis that states that a child's target game performance is influenced by the racial composition of the group in front of which he is performing at that time would lead to exactly the same prediction for this data.

What are the possible theoretical interpretations of the effects so far presented? The effects of sex (females lower) and race (Whites lower) affirm the preliminary findings reported in FitzGibbon (1971) and are not inconsistant with the hypothesis that low discrepancy scores indicate 'healthy self concept.' The effect of grade (second graders lower) is better interpreted when the current scoring system is more closely examined.

Assume for the moment that second graders choose to call within a closer larger box target area because it is realistic for them to do so, and that third graders choose to call within a farther large box target area because it is realistic for them to do so. A known bias in the current scoring method assigns greater scoring method discrepancies to farther target areas whenever any pair of large box, any circle, and "off" the targets are compared.

The conclusion is that higher scores may be more realistic for third graders and lower scores may be more realistic for second graders. If this is correct, and if the hypothesis relating sex effect to 'healthy self concept' is also correct, then the current scoring method may be indexing two facets which can work in opposite directions at times. If then, we are to substantively interpret differences among districts and differences among schools in the same district, an intensive analysis of the internal structure of the target game scoring is necessary; exploration of possible alternative methods of scoring will be investigated.

The statistical analyses so far have used linear models which more or less presuppose some things about distribution, additivity, and scale. Analyses which follow will examine these characteristics at the cost of ignoring between school effects which are known to be salient.

Although the incremental effect of race is lost when school boundaries are ignored, further analysis showed that the effect of grade is not uniform

with respect to either race or different parts of the target game scale.

Uniformity of differences in Target Game scores with respect to race, grade, and sex are shown in Figures 2-5. Each of these figures is a probability plot and can be interpreted as follows:

1. Target Game score appears on the horizontal scale given at the top and bottom of the figure. Percentile, or percent of children scoring less, appears on the vertical scale at the right of the figure.

2. High Scores are on the right. Thus, if two curves do not cross, the curve on the right has higher scores at every percentile.

3. Slope is related to concentration of scores. A steep curve represents a concentrated distribution or small standard deviation; a shallow curve represents a spread out distribution or large standard deviation.

4. Straightness is related to 'normality' of distribution. An approximately straight line represents an approximately normal (Guassian) distribution. The manner in which a curve departs from straightness represents the manner in which the distribution is non-normal:

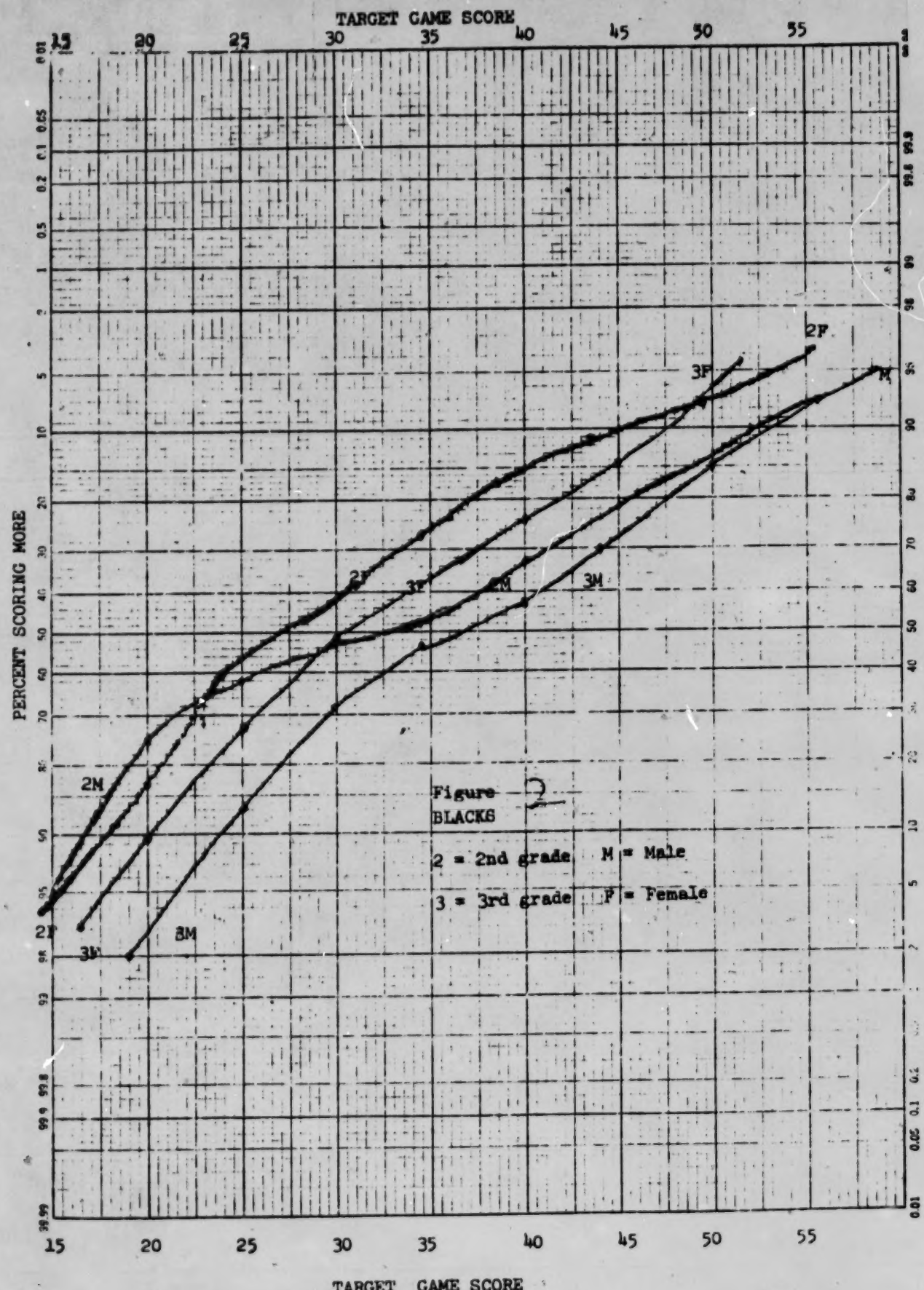
   a) Too flat on the right: high scores are extra high

   b) Too flat on the left; low scores are extra low

   c) Turned up on the right: ceiling effect

   d) Turned down on the left: floor effect

The smoothed curves of Figures 2-5 are based on the following group sizes:


TARGET GAME SCORE
15 20 25 30 35 40 45 50 55
PERCENT SCORING MORE
PERCENT SCORING LESS
2F
3F
2M
3M
2F
3F
2M
3M
2M
2F
3F
3M
Figure 2
BLACKS
2 = 2nd grade M = Male
3 = 3rd grade F = Female
TARGET GAME SCORE

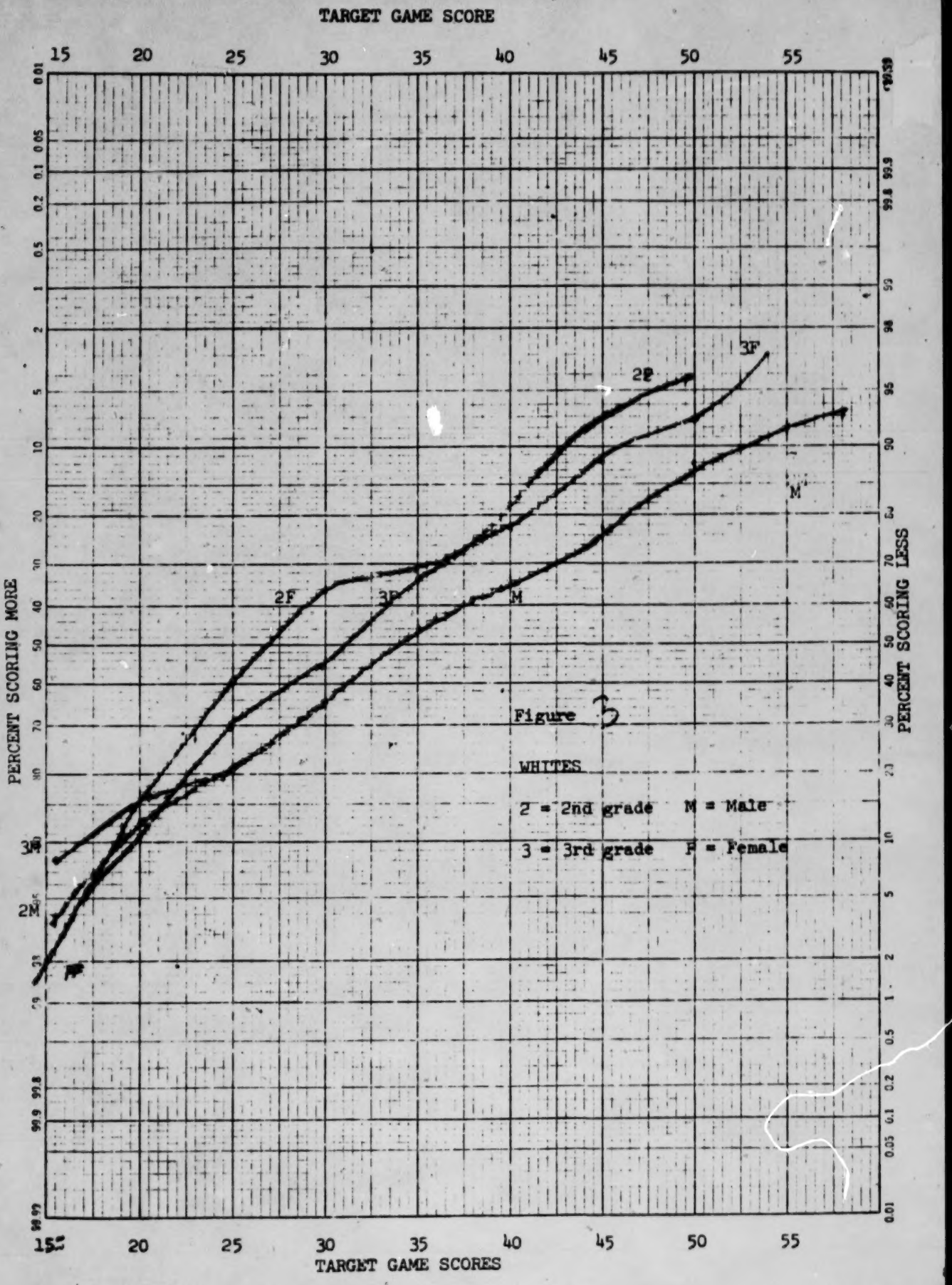
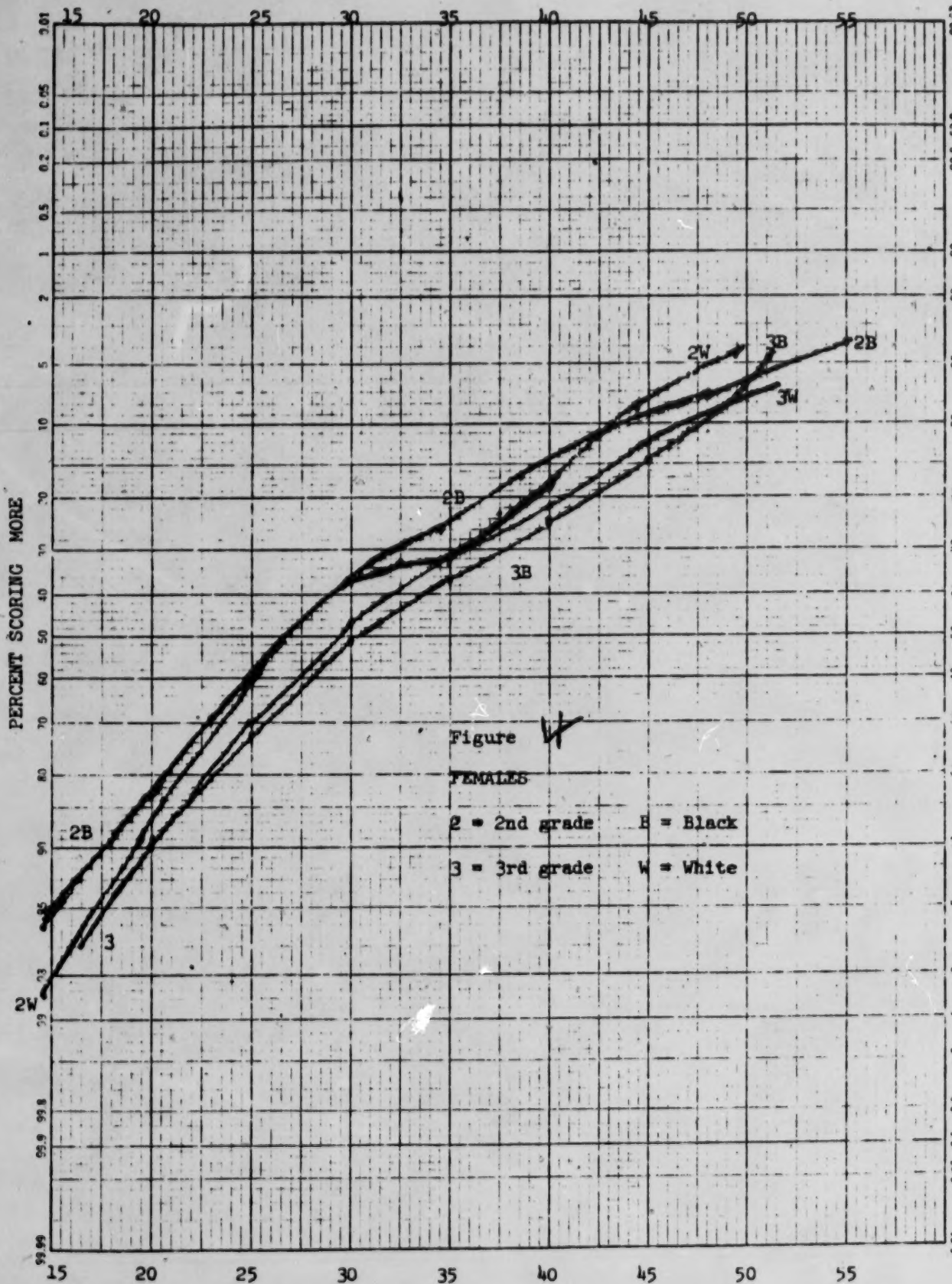


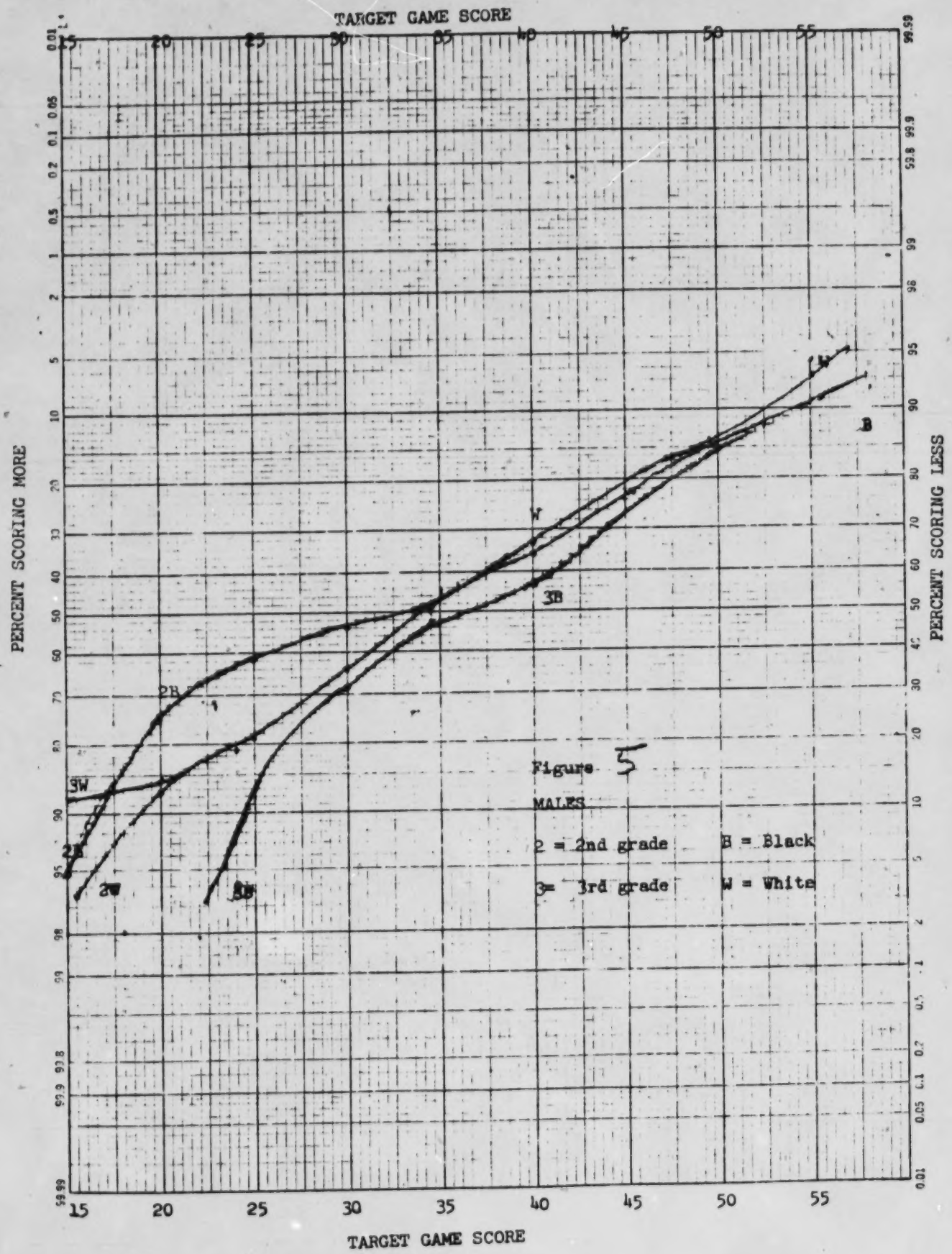
Figure 5

WHITES

2 = 2nd grade M = Male

3 = 3rd grade F = Female


PERCENT SCORING MORE
PERCENT SCORING LESS
15
20
25
30
35
40
45
50
55
2W
3B
2B
3W
2B
3B
2B
3
2W
Figure
FEMALES
2 = 2nd grade
B = Black
3 = 3rd grade
W = White



K-E PROBABILITY X 90 DIVISIONS KEUFFEL & ESSER CO.

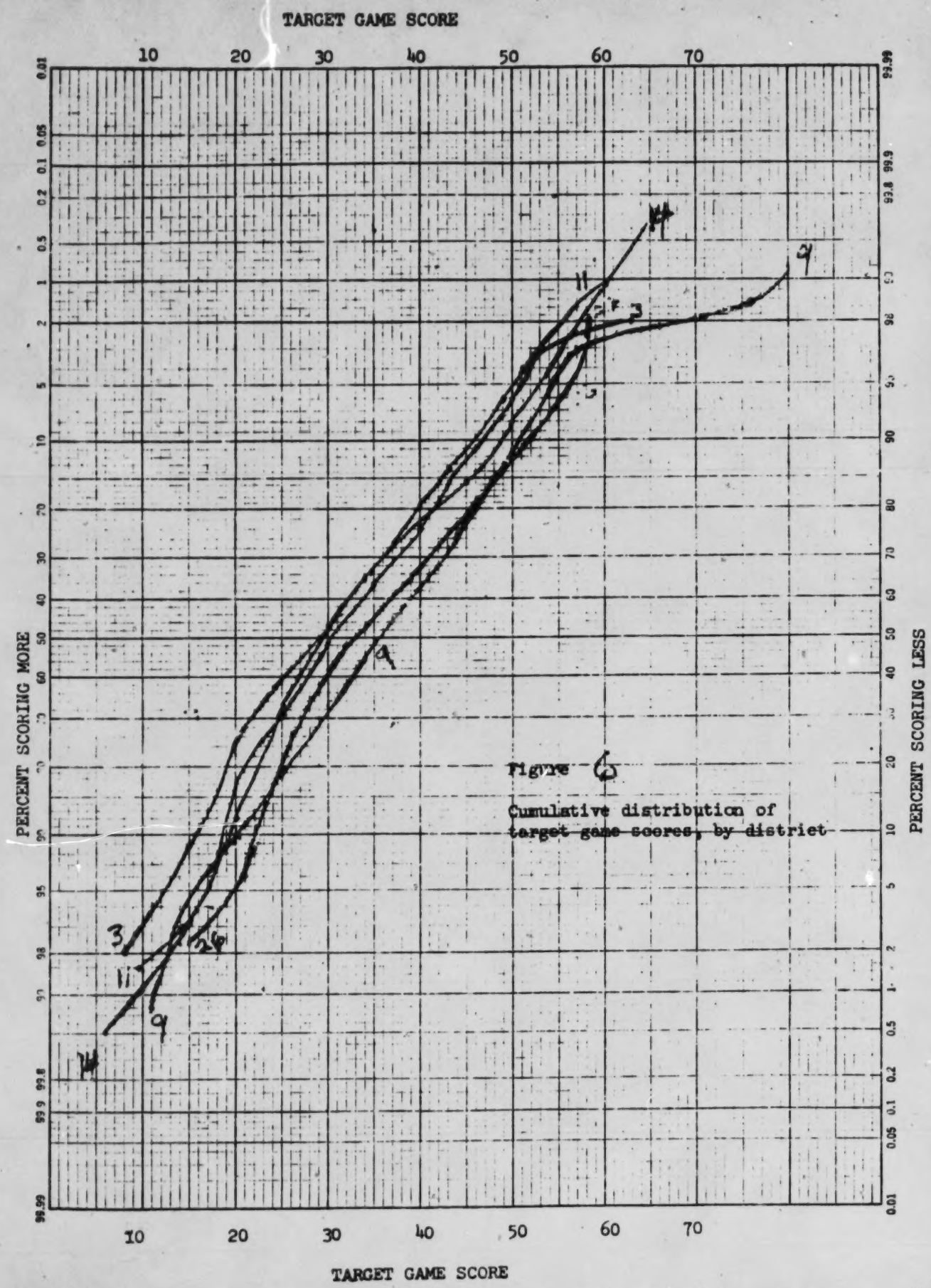


Figure 6

Cumulative distribution of target game scores, by district

	2nd grade	3rd grade
Male-Black	57	52
Female-Black	52	61
Male-White	55	72
Female-White	65	65

Figure 2 shows the effects of grade and sex for Blacks. For higher ranging scores between the 50th and 90th percentiles, both male curves are to the right of both female curves: males score higher (are more discrepant) than females independently of grade. For each sex 3rd grade curves appear to the right of corresponding 2nd graders. For lower ranging scores between the 10th and 50th percentiles, however, this pattern is changed in that the curve for 2nd grade males crosses the curves for females: low scoring 2nd grade males are extra low, thereby neutralizing the effect of sex for low-scoring 2nd grade Blacks.

Figure 3 shows the effects of grade and sex for Whites. For scores above the 20th percentile, both male curves are to the right of both female curves. Males score higher than females, independently of grade. For females the 3rd graders score higher than 2nd graders, but the effect of grade is neutralized near the 72nd percentile for female Whites. The 2nd and 3rd grade curves for males virtually coincide: the effect of grade is neutralized for Male Whites.

The same curves are redisplayed in Figures 4-5. Figure 4 shows the effects of grade and race for females. Between the 10th and 90th percentiles, both curves for 3rd graders appear to the right of both curves for 2nd graders; 3rd graders score higher than 2nd graders independently of race. The curves for Blacks 'surround' the curves for Whites: the effect of grade is

attenuated for White Females.

Figure 5 shows the effects of grade and race for males. For high scores above the 60th percentile the curves are close together, i.e. among high scoring males, the effect of grade and race are slight, but the curves for Blacks appear to the right, with the difference virtually neutralized in the 2nd grade. Among lower scores, the curves for Blacks 'surround' those for Whites, i.e. the effect of grade is neutralized for White Males.

The main results of these observations on the effects of sex, grade, and race, ignoring school boundaries, may be summarized as follows: The effects of sex and grade are incremental; the effect of race is operational. Males score higher than females, but low-scoring 2nd grade male Blacks are extra low, neutralizing the effect of sex for 2nd grade Blacks. Third graders score higher than 2nd graders, but the effect of grade is attenuated for Whites and neutralized for White males.

Figure 6 shows smoothed percentile graphs for each district. Within the middle 80% of the distribution, Berkeley; Lebanon and Owensboro have substantially similar curves, while the curve for Goldsboro appears on the right indicating higher scores for that district. The curves for Omaha is surrounded by these for the middle 60% of the distribution.

3. Characteristics of Target Game Items

Item means, standard deviations, and correlations between them and with total scores are among the common statistics reported in test manuals. This section reports these characteristics, both pooled and separately for each school district.

Item means and mean Target Game Scores for each district are shown in Table 9. Means do not differ substantially from item to item or across districts. Table 10 shows standard deviations of items and total scores for each district and for the total group. The standard deviations of target game scores are substantially similar across districts, and it is reasonable to consider the pooled value, 11.6, as representative.

Table 11 reports item by item and item by total score correlations, item means and standard deviations for the total sample of 553 subjects. Similar tables computed for each district separately can be found in Appendix A (Tables 16-20).

4. Reliability of the Target Game

The item by subject Analysis of Variance is shown in Table 12; reliability estimates resulting are shown in the last column. The entries under 'subject mean square' (test score variance divided by number of items) are substantially similar across districts, so that the pooled result represents all. The entries in the 'within mean square' estimate that portion of the 'subject mean square' with respect to which the items do not agree that the subjects are different; these too are substantially similar across districts. The reliability coefficient estimates are also substantially similar across districts. Using the pooled estimates, the following quantities may be computed:

TABLE 9

Target Game Item Means by District and Pooled.

District	Berkeley	Goldsboro	Owensboro	Lebanon	Omaha	pooled
n item score	50	129	260	65	41	553
1	3.36	3.44	3.08	2.65	3.00	3.15
2	3.24	3.97	3.48	3.43	3.32	3.55
3	3.12	3.44	3.19	3.06	4.00	3.29
4	3.24	3.67	3.43	3.26	4.07	3.49
5	2.52	3.52	3.03	3.15	3.22	3.13
6	3.48	3.72	3.19	3.55	3.76	3.41
7	3.54	3.71	3.31	3.17	3.54	3.44
8	1.96	3.64	3.12	3.00	3.41	3.15
9	3.10	3.53	3.18	3.46	3.41	3.31
10	2.54	3.39	3.08	3.31	3.17	3.14
Mean target game score	30.10	36.03	32.09	32.00	32.00	33.06

TABLE 10

Target Game Item Standard Deviation
by District and Pooled.

District	Berkeley	Goldsboro	Owensboro	Lebanon	Omaha	Pooled
item score	50	129	260	65	41	553
1	2.71	3.08	2.53	2.03	1.88	2.60
2	2.60	2.94	2.65	2.74	2.60	2.72
3	3.03	2.78	2.49	2.50	2.33	2.60
4	2.51	2.86	2.70	2.62	2.95	2.73
5	2.36	3.17	2.47	2.70	2.39	2.67
6	2.66	3.28	2.66	2.64	2.75	2.81
7	2.73	3.02	2.82	2.57	2.56	2.80
8	2.17	3.14	2.54	2.35	2.66	2.69
9	2.73	2.88	2.53	2.56	2.41	2.64
10	2.64	2.88	2.68	2.82	2.51	2.72
target game score	11.53	12.27	11.18	10.98	10.89	11.56

TABLE 11

Target Game. Item by Item by Total Correlations and Item Means Standard Deviations for 5 districts pooled (n = 553)

item	1	2	3	4	5	6	7	8	9	10	target game
1	1.00	.12	.08	.15	.17	.14	.13	.14	.07	.07	.476
2	.12	1.00	.15	.04	.14	.07	.01	.10	.04	.08	.416
3	.08	.15	1.00	.09	.11	.04	.07	.09	.13	.08	.419
4	.15	.04	.09	1.00	.17	.04	.06	.15	.14	.04	.439
5	.17	.14	.11	.17	1.00	.10	.06	.10	.12	.08	.478
6	.14	.07	.04	.04	.10	1.00	.07	.06	.04	.06	.389
7	.13	.01	.07	.06	.06	.07	1.00	.09	.06	.02	.373
8	.14	.10	.09	.15	.10	.06	.09	1.00	.15	.12	.466
9	.07	.04	.13	.14	.12	.04	.06	.15	1.00	.11	.428
10	.07	.08	.08	.04	.08	.06	.02	.12	.11	1.00	.391
target game	.476	.416	.419	.439	.478	.389	.373	.466	.428	.391	
means	3.15	3.55	3.29	3.49	3.13	3.41	3.44	3.15	3.31	3.14	33.06
SD	2.60	2.72	2.60	2.73	2.67	2.81	2.80	2.69	2.64	2.72	11.56

TABLE 12

Item by Subject Analysis of Variance, Mean Squares and Reliability Coefficient Estimates.

district	n	item mean square	subject mean square	within mean square	reliability estimate
Berkeley	50	12.85	13.30	6.16	.54
Goldsboro	129	3.94	15.06	8.37	.46
Owensboro	260	6.01	12.39	6.18	.50
Omaha	41	5.14	11.86	5.74	.52
Lebanon	65	4.67	12.05	5.96	.51
pooled	553	14.15	13.30	6.62	.50

Reliabilities are significantly different from zero at levels $< .001$.

Statistic	Value	Formula
Reliability coefficient	.50	$\rho = 1 - \frac{\text{within MS}}{\text{subject MS}}$
Coefficient of alienation	.71	$k = \sqrt{1-\rho}$
Standard error of target game score	8.1	$\sqrt{10 \times (\text{within MS})}$
Standard error of the target game Percentiles near the median	.26	$\Phi(\kappa) - 1/2$ where Φ is the cumulative standard normal distribution function.

When the reliability estimates in Table 12 are compared with the .87 obtained in the second field test, these new estimates are consequently lower and seem discouraging.

With respect to their size, it is reasonable to use reliability coefficients as a basis of comparison of instruments which are otherwise almost equal. However, consider first a 'standard' style spelling test in which the subject is presented with a fixed list of words, asked to spell them and scored. Internal consistency analysis reveals the extent to which the items agree on how the subjects are different. Now consider another type spelling test in which the subject is asked to name ten words and asked to spell them, and scored. Internal consistency analysis again reveals the extent to which the items agree that the subjects are different, but the items are not the same for different subjects. It is manifestly unreasonable to compare two such tests on the basis of reality coefficients. Now suppose, contrary to fact, that memory is not a factor, and the tests are re-administered. For the 'standard' test, we would expect test-retest analysis to give almost the same picture as 'internal consistency analysis.' For the other style test, if the subject decides to name substantially the same list of words and his ability to spell them has not changed much, we would expect test-retest coefficients

to be high, even though internal consistancy coefficients were modest. If repeatability is all one cares about, then perhaps it is reasonable to compare test-retest coefficients of the different instruments.

The target game is analogous to the second style test in the sense that the player creates a situation and is then scored relative to that situation; but on any trial, different players are in different situations. If discrepancy scoring in the target game is to be compared to other instruments on the basis of internal consistancy reliability, then this comparison should be limited to instruments of the same type. For purposes of standard errors of test scores, however, internal consistency analysis may well be more appropriate than test-retest in the sense of already taking into account that scores obtained by different subjects are obtained in different circumstances. Suitable decomposition of discrepancy scores into their component parts may well render this question academic as far as the target game is concerned. Devoting considerable research attention to improving the reliability of a test is not necessarily more virtuous than using it as is. The cost of improvement may far exceed the value of the gain obtained. Psychometric theory tells us that a higher reliability instrument will be adquate in a wider variety of circumstances than a lower reliability instrument, other things being equal. What this theory does <u>not</u> do is offer magic numbers above which reliability is 'good' and below which it is 'bad'.

Although the reliability coefficient, the coefficient of alienation, and the standard error of percentile estimates are all mathematically interrelated, it is the coefficient of alienation (i.e. what the standard error would be if the standard deviation of the instrument were fixed up to be one) that is most pertinent to individual assessment problems. Percentile

representations are widely used in education, so looking at standard errors of percentiles near the median can be helpful. The standard error of percentiles is used in the construction of graphs such as that of Figure 1 of this report; it is the vertical distance between the dotted curves and the solid curve of such figures. The values of some of these statistics for some reliabilities is given in Table 13. For children about whom one knows nothing, one can guess that their underlying percentiles are within the limits 50% $\pm$ 34% and be right sixty-eight percent of the time, so the standard error of percentiles near the median for a worthless insturments is .34. For the Target Game Score, this is reduced to .26. Table 13 makes it clear that one needs substantial gains on the reliability scale to achieve modest cuts in standard errors of individual scores or percentiles. (e.g. to cut current standard errors in half would require boosting the reliability to .90 under current conditions.) Psychometric theory also offers a hypothetical projection of the number of items needed to attain a particular reliability provided it could be done without changing the nature of the instrument. Exemplary projections are also displayed in Table 13. It cannot be too strongly emphasized that the assumptions underlying such projections are often not met in practice and can fail to deliver the increased accuracy promised. If the projections were right, which is unlikely, it would take a nine-fold increase in the number of items to cut the current standard error of percentiles near the median in half. While this is not the road to take in general, there is one situation for which such projections are pertinent: if a researcher is interested in a serious clinical appraisal of how the target game works for an individual child, a substantial sample of that child's target game behavior may be needed.

TABLE 13

Various Statistics Related to Reliability

reliability coefficient	coefficient of alienation	standard error of percentiles near the median	projected number of items needed under the assumption that the instrument is not changed thereby
0.	1.	.34	0
.25	.87	.30	3
.50	.71	.26	10
.70	.55	.21	23
.80	.45	.17	40
.90	.32	.13	90
.95	.22	.09	190
.99	.10	.04	990
1.00	0.00	0	infinite

TABLE 14

Projected Needed Sample Sizes to Increase Reliability

desired standard error of difference	size of smaller group*
.50	1064
.75	473
1.00	266
1.25	170
1.50	118
1.75	87
2.00	67

*Discount by 25% if the larger group is twice as large; discount by 50% if the larger group is many times as large.

For research purposes in which the goal is to relate target game scores to external variables, the interpretation of reliability is straight-forward when the relations are modest: the reliability coefficient times the sample size used is the size of the sample that would have done as well for an instrument with reliability of 1.00. The implication of the target game reliability coefficient of .5 is that you need twice as many cases to achieve the same precision in estimating modest group differences as would be needed for an instrument with reliability close to 1.00. If a sample size of 300 is needed to attain a certain precision for a test with reliability of .5, a test with reliability of.75 would get the same precision for 200 cases; according to the item prediction formula, a three-fold increase in the number of items would be necessary to achieve this 33% reduction in sample size. Approximate sample sizes needed for the simplest two group comparison situations are displayed in Table 14.

If the standard error of the difference between two groups means is as low as 1.5, as would be appropriate for detecting statistical significance of a difference of 3 points on the target game scale then on needs about 118 cases in each group.

On the basis of the considerations outlined above, the recommendations of the statistician and for this study concerning the amount of research to be directed toward improving the reliability of Target Game Scores as the instrument is currently scored and administered comes was none.* It is sufficient to live with current reliability levels and use the score in research contexts where adequate group sizes are available.

*The author is grateful to David Piezer for his valuable suggestions and critisms made during his statistical analysis of the Target Game data.

5. Validity of Target Game Scores

During the summer of 1972, a study was undertaken using the Self Concept Target Game and a Cognitive Skills Game, also measuring level of aspiration. The data obtained from the Cognitive Skills Game indicated that it had too low a ceiling and that individual questions were too similar in difficulty. Children who got all the items right on the demonstration trial knew just how they would do on the remaining trials, and were usually right, and children who missed items were less strongly impressed with their previous results.

It was hoped that information on the two tests administered to a sample in Berkeley would give us valuable validity information. At present, there no other known instrument having a body of data resources level of aspiration in young children. This instrument does have face validity; in the two earlier field tests predictive validity was evidenced since the Target Game distinguished children enrolled in a Responsive Model Follow Through Program. However, in this field test, being in a Responsive Model Follow Through program does not seem to make a difference in Target Game Scores.

6. Some Problems Noted in Connection with the Interpretation of Test Results.

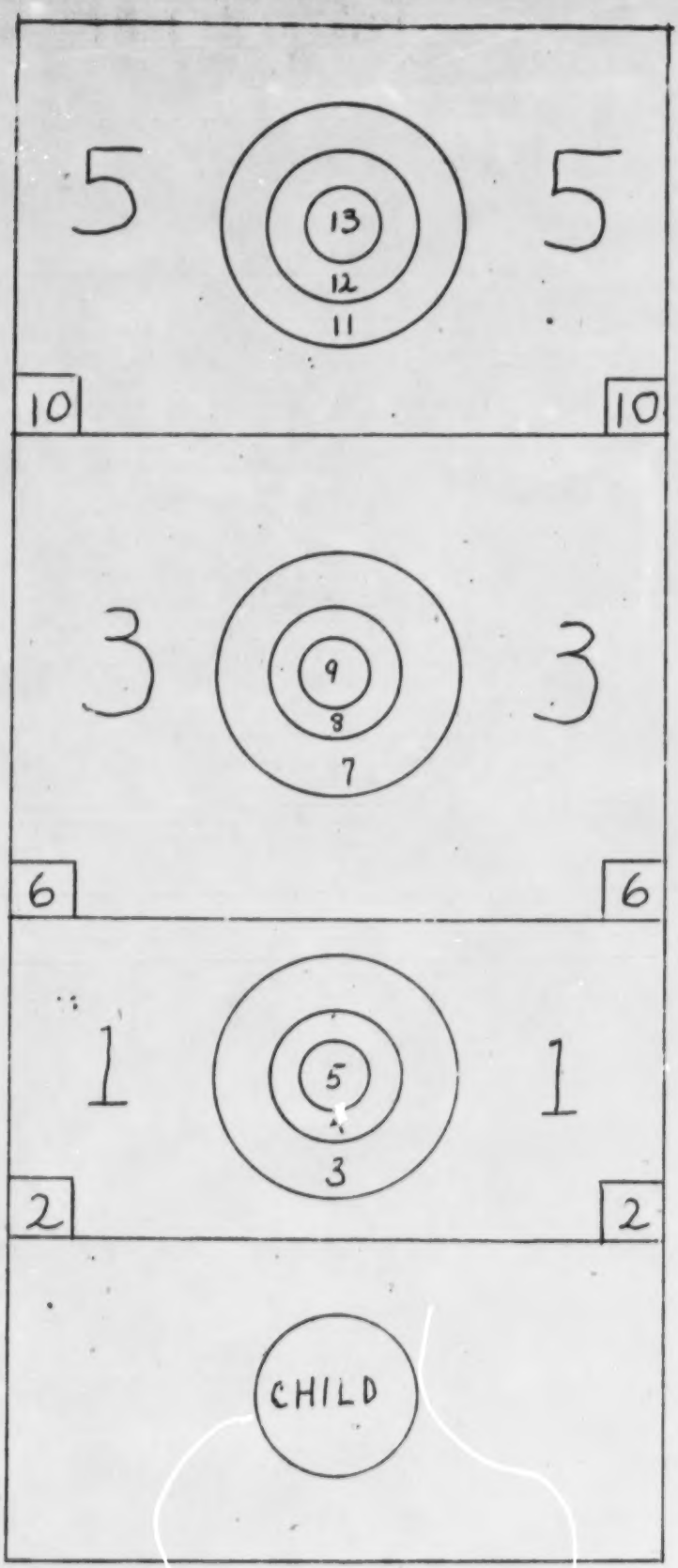
a) Group Testing.

Open group testing, while convenient at times, considerably complicates both the analysis and interpretation of data. In the case of the Target Game and the research reported here, open group testing confounds examiner effects and audience effects with all of the other effects under consideration. It has been shown here that Target Game Score averages are lower

for females than for males. Perhaps, instead of females actually scoring lower than males, a competing interpretation is that the scores of children of both sexes are lower in front of more predominantly female audiences than in front of more predominantly male audiences; since the children in the audience and the children tested are the same for each group, and since there is natural variation in the proportion of children of either sex from group to group, the two possible sources of variation, individual and audience, are confounded. In general, audience effects include kibitzing, imitating previous children, trying to be different from previous children, and trying to do better than previous children. The study of audience effects might prove interesting, but the goals of the current research seem to indicate that individual testing is more desirable. In further administrations of the Target Game, it is therefore planned to administer individually to the subjects. The amount of time lost will be, it is felt, more than compensated for by the gain in accuracy of the information obtained.

b) Physical Format of The Target Game and Resulting Scoring.

Figure 7 shows schematically the physical format of the Target Game; numbers on the target enter into both the play and the scoring directly. The discrepancy scoring method uses the same numbers as are shown on the Target. Since using the same values for both play and scoring can produce certain oddities, this points to using other numbers for scoring than are used in the Game. Administrators noted that children missed the strip completely at times, and sometimes called boxes after they hit circles in the same area. A scoring oddity occurs whenever a comparison is made between any two 'off the target', large box, and any circle; discrepancy scores reckoned by the present scoring system are smaller for near target areas, and larger for farther target areas.



Numbers in the diagram are numbers on the target

Figure 7.
THE TARGET

With the present scoring, the correlations among calls and results can enter into the trial-to-trial internal consistency of the discrepancy scores, and a highly tentative analysis, based only on the Berkeley data, suggests that this is the case here. Table 15 shows the intercorrelations, means and standard deviations of the calls and results for each of the eleven trials in that district. Examining the means, it can be seen that calls exceed previous results by an average of one or two points, so overestimates predeominate over underestimates on each trial. Note that almost every call is more highly correlated with the result on the same trial, and vice versa, than with anything else. If this, and the most positive correlations were all that were known, one would predict that the correlations among the discrepancy scores might be negative, especially so for successive differences. The low correlations elsewhere, and their consequences for internal consistency, may be tentatively viewed as a result between what might be theoretically attainable, and its reduction due to the combination of imbalance of over/underestimates and high correlation of call and result on the same trial.

While modifications of the physical layout of the target may turn out to desirable, it seems best as present to look at other methods of scoring, and what other types of information are available from the data on the score sheets. The simple discrepancy score is only part of the information potentially available here. In general, we need to know:

a) within box calls, results, overestimates, and discrepancies.

b) between box calls, results, overestimates, and discrepancies.

c) full target calls, results, overestimates, and discrepancies.

Table 15

	1c	1r	2c	2r	3c	3r	4c	4r	5c	5r	6c	6r	7c	7r	8c	8r	9c	9r	10c	10r	11c	11r
1c		.455	.425	.229	.347	.247	.352	.101	.156	.144	.240	.006	.277	.037	.209	.212	.304	.218	.247	.146	.278	.114
1r	.455		.219	.214	.252	.247	.290	.137	.188	.173	.038	-.025	.128	.023	.156	.239	.128	.131	.315	.087	.226	.093
2c	.425	.219		.544	.289	.262	.332	.153	.384	.367	.367	.052	.374	.209	.276	.288	.433	.298	.457	.251	.352	.242
2r	.229	.214	.544		.302	.144	.328	.268	.153	.188	.077	.086	.247	.100	.229	.287	.175	.193	.276	.170	.332	.176
3c	.347	.252	.289	.302		.629	.454	.155	.151	.175	.337	.331	.365	.132	.226	.014	.154	-.050	.349	.274	.525	.073
3r	.247	.247	.262	.144	.629		.286	.138	.060	.043	.084	.101	.119	.023	.069	-.015	.055	-.098	.098	-.088	.252	.285
4c	.352	.290	.332	.328	.454	.286		.508	.313	.337	.375	.172	.281	.012	.398	.286	.323	.040	.476	.362	.505	.215
4r	.101	.137	.153	.268	.155	.138	.508		.371	.420	.203	.019	.285	.170	.170	.173	.072	.035	.202	.265	.283	.183
5c	.156	.188	.384	.153	.151	.060	.313	.371		.790	.576	.202	.347	.245	.279	.243	.500	.362	.484	.435	.378	.261
5r	.144	.173	.367	.188	.175	.043	.337	.420	.790		.528	.115	.310	.275	.088	.014	.412	.244	.376	.278	.306	.264
6c	.240	.038	.367	.077	.337	.084	.375	.203	.578	.528		.410	.590	.464	.240	.087	.586	.393	.437	.333	.265	.009
6r	.006	-.025	.052	.086	.334	.101	.172	.019	.202	.115	.410		.347	.249	.150	-.045	.254	.187	.173	.165	.232	.053
7c	.277	.128	.374	.247	.365	.119	.281	.285	.347	.310	.590	.347		.676	.390	.249	.566	.385	.374	.323	.382	.078
7r	.037	.023	.209	.100	.132	.023	.012	.170	.245	.275	.464	.249	.676		.085	-.021	.256	.213	.240	.145	.112	.074
8c	.209	.156	.276	.229	.226	.069	.398	.170	.279	.088	.240	.150	.390	.085		.694	.533	.341	.364	.194	.248	.079
8r	.212	.239	.288	.287	.014	-.015	.286	.173	.243	.014	.087	-.045	.249	-.021	.694		.542	.488	.284	.164	.320	.212
9c	.304	.128	.433	.175	.154	.055	.323	.072	.500	.412	.556	.254	.566	.256	.533	.542		.705	.264	.211	.255	.151
9r	.218	.131	.296	.193	-.050	-.098	.040	.035	.362	.244	.393	.187	.385	.213	.341	.488	.705		.156	.049	.176	.126
10c	.247	.315	.457	.276	.349	.098	.476	.202	.484	.376	.437	.173	.374	.240	.364	.284	.264	.156		.745	.530	.133
10r	.146	.087	.251	.170	.274	-.088	.362	.265	.435	.278	.333	.165	.323	.145	.194	.164	.211	.049	.745		.475	-.110
11c	.278	.226	.352	.332	.525	.252	.505	.283	.378	.306	.265	.232	.382	.112	.248	.320	.255	.176	.530	.475		.527
11r	.114	.093	.242	.176	.073	.285	.215	.183	.261	.264	.009	.053	.078	.074	.079	.212	.151	.128	.133	-.110	.527	
mean	5.67	3.89	5.90	4.48	5.38	3.87	5.98	3.95	6.07	5.03	6.18	3.90	6.11	4.44	6.00	4.34	5.48	4.39	5.74	4.54	5.77	4.02
s. d.	3.34	2.80	3.54	2.94	3.80	3.23	3.34	2.97	3.34	3.04	3.66	2.58	3.56	3.09	3.37	2.85	3.37	3.26	3.20	2.91	3.67	2.83

Correlations, Means, and Standard Deviations of Calls and Results on Eleven Trials of the Target Game (n=61)

d) the game the child plays: number of trials rewarded, total reward.

e) unfavorable choice: frequency of two, six or ten box call.

f) relation of call to whether the previous hit was rewarded (up after success, down after failure)

Extensive internal analysis of the data already available is a prerequisite to determining what aspects of the current physical layout can or should be changed. Suggestions for such an analysis can be found in Appendix B.

Summary

The Self Concept Target Game was designed to measure the Level of Aspiration of second and third grade children; a review of the literature indicated that a realistic Level of Aspiration may be one of the factors in a healthy self concept. Since promoting the growth of a healthy self concept is one of the goals of Far West Laboratory's Responsive Model Head Start and follow Through programs, a series of tasks are being designed to measure various aspects of self concept, instead of attacking the problem in a global manner.

The Self Concept Target Game has undergone three field tests with second and third graders throughout the country. In the first field test, a simple procedure of sorting the test protocols into those who made realistic choices of targets (up after success and down after failure) indicated that our game had some validity since the results correlated with teachers' choices of children high and low in a healthy self concept.

In scoring the results of the second and third field test, a simple discrepancy score was used as a measure of performance, i.e., the differences between the hit on one trial and the choice on the next trial were summed, ignoring the differences. Results of the second field test indicated that Responsive Model Follow Through children performed better (had lower discrepancy scores) than children not exposed to the special program, and that the most realistic performance was given by children whose teachers scored highest on a Responsive Observation Schedule for teachers, and who were black. While white children out-performed Blacks, Black children in integrated classrooms did better than Black children in segregated classrooms; in general, girls did better than boys. Reliability for the test was found to be .87.

The third and largest field test included 553 children second and third

graders in 5 shool districts throughout the country; the large sample allowed for determining various characteristics of the Target Game. The number of trials on the Game was increased from 6 to 11 in the hope of increasing the reliability. Results indicated that of the 10 possible comparisons were significantly different. In general, second graders did better than third graders; blacks were less realistic than whites; boys less than girls, and enrollment in the Responsive Model Follow Through program did not seen to make a difference in score. While these differences appeared for the total group and each school district as well, when school boundaries are ignored, difference between grades is reduced, and differences between races disappear. Schools in the same district differed more from one another than did one district from another. In addition to reporting the means and standard deviations of the total scores, this report also contains the means and standard deviations of individual item scores, and the correlations among the items and with total score, separately for each school district and for the total group. Reliability for the total test in the 5 districts ranged from .46 to .54. Two problems were noted in connection with the Self Concept Target Game, first, the one connected with group testing, and the second with the present discrepancy scoring system, since discrepancies are smaller for near target areas and larger for farther target areas. There is also much potential information available from the score sheets, including pattern of choice and number of realistic choices, that have not been taken into consideration in using the discrepancy score. No information is available with regard to the relation of the Self Concept Target Game to other tests of self concept or levels of aspiration. The instrument can be used by other researchers with the permission of Far West Laboratory, providing that credit is given the Laboratory, and that data is made available to us.

REFERENCES

1. Birney, R.C., Burdick, H. and Teevan, R.C. Fear of Failure. New York Van Nostrand, 1969.

2. Borislow, B. Self evaluation and academic performance. J. counsel. Psychol., 9, 1962, 246-254.

3. Brim, O., Glass, D., Neulinger, J. and Firestone, I. American beliefs and attitudes about intelligence. New York: Russell Sage Foundation, 1969.

4. Fitzgibbon, A. Self-Concept target game. Revised report. Berkeley: Far West Laboratory for Educational Research and Development, Dec. 1971.

5. Frank, J.D. Individual differences in certain aspects of the level of aspiration. Amer. J. Psychol., 1935, 47, 119-128.

6. Goldblott, H. and Tyson, C. Some self perceptions and teacher evaluations of Puerto Rican, Negro and White pupils in 4th, 5th and 6th grades. ERIC ED001776, 1962, 25p.

7. Guggenheim, F. Self esteem and achievement expectations for White and Negro children. New York City Board of Education, ERIC ED021938, 1967.

8. Harrison, R. etal. Level of aspiration as a function of learning potential status in the educable mentally retarded. Vol. I. Cambridge: Research Institute of Educational Problems. ERIC ED048706, 1971.

9. Kaufman, H. Task performance and responses to failure as functions of imbalance in the self concept. Psych. Monogr, 77 (6) whole no. 569, 1968, 1-15.

10. Lasky, C. Level of aspiration and locus of control in disadvantaged children. Psych. Rep, 1970, 27, 343-350.

11. Lefcourt, H. Lewis, L. and Silverman. T. Internal vs. external control of reenforcement and attention in a decision making task. J. Pers., 36, 1968, 665-682.

12. Mitchell, J.V. Goal setting behavior as a function of self acceptance, over and underachievement, and related personality variables. J. educ. Psychol., 50, 1959, 93-104.

13. Mukherjee, B. Self perceptions and Achievement Motivation. J. Psych. Res. 1965, 9 (3), 109-114.

14. Nimnicht, Glen P. A revision of the basic program plan of education at age 3. Berkeley: Far West Lab, April, 1970.

15. Redfearn, T. Level of expectation, actual performance, and reactions to success and failure in 3 ethnic groups. Paper presented at the Amer. Psychol. Assn. Meeting, Washington, 1969.

16. Sears, Pauline S. Levels of aspiration in academically successful and unsuccessful children. J. abn. Soc. Psych., 1940, 35, 498-536.

17. Trowbridge, N. Self concept and socioeconomic status in elementary school children. AERA J., Fall, 1972, 9 (4), 525-537.

APPENDIX A

Tables 16-20 report item by item and by total score correlations for each district along with item and tota and standard deviations previously presented. It is to be noted th arate district reports of item by item correlations are given here for accounting purposes only: the size of the typical correlation under estimation is about 1/11. One needs about 125 cases to achieve a standard error associated with estimating such a correlation as small as 1/11; one needs about 500 cases to reduce the standard error of estimate to 1/22 i.e. half the size of the typical quantity under estimation.

Thus, although Tables 16-20 display some of the basic characteristics that would be reported in a test manual, and displays these separately for each district, the quantities estimated are not in focus. A much more focused picture is given by Table 14.

TABLE 16

Target Game Item by Item and
Item by Total Correlations,
Item and Total Means and
Standard Deviations for (n=50)

Berkeley

item	1	2	3	4	5	6	7	8	9	10	total
1	1.00	.25	.02	.16	.26	.36	.15	.37	-.03	-.01	.563
2	.25	1.00	.02	.15	.12	.37	.22	.18	.00	-.11	.490
3	.02	.02	1.00	.09	.22	.12	.11	.02	.29	-.27	.403
4	.16	.15	.09	1.00	.15	-.02	-.06	.32	.23	.06	.453
5	.26	.12	.22	.15	1.00	.28	.23	-.20	.30	-.09	.516
6	.36	.37	.12	-.02	.28	1.00	.37	.16	-.04	-.08	.575
7	.15	.22	.11	-.06	.23	.37	1.00	.04	-.07	-.12	.431
8	.37	.18	.02	.32	-.20	.16	.04	1.00	.15	-.03	.425
9	-.03	.00	.29	.23	.30	-.04	-.07	.15	1.00	.15	.455
10	-.01	-.11	-.27	.06	-.09	-.08	-.12	-.03	.15	1.00	.110
total	.563	.490	.403	.453	.516	.575	.431	.425	.455	.110	
mean	3.36	3.24	3.12	3.24	2.52	3.48	3.54	1.96	3.10	2.54	30.10
SD	2.71	2.60	3.03	2.51	2.36	2.66	2.73	2.17	2.73	2.64	

TABLE 17

Target Game Item by Item and
Item by Total Correlations, Item
and Total Means and Standard Deviations
for (n=55)

Lebanon, N. H.

item	1	2	3	4	5	6	7	8	9	10	T
1	1.00	.03	-.10	-.03	.10	.01	.08	-.03	-.11	.21	.227
2	.03	1.00	.17	-.09	.01	-.08	-.05	.16	.24	.17	.375
3	-.10	.17	1.00	-.02	.17	.08	-.03	-.01	.26	.32	.444
4	-.03	-.09	-.02	1.00	.00	.02	.13	.08	.13	.20	.340
5	.10	.01	.17	.00	1.00	-.02	.12	.24	.24	.14	.474
6	.01	-.08	.08	.02	-.02	1.00	.01	.10	.02	.12	.299
7	.08	-.05	-.03	.13	.12	.01	1.00	.24	.13	-.06	.358
8	-.03	.16	-.01	.08	.24	.10	.24	1.00	.43	.11	.533
9	-.11	.24	.26	.13	.24	.02	.13	.43	1.00	.26	.616
10	.21	.17	.32	.20	.14	.12	-.06	.11	.26	1.00	.591
total	.227	.375	.444	.340	.474	.229	.356	.533	.616	.591	
mean	2.65	3.43	3.06	3.26	3.15	3.55	3.17	3.00	3.46	3.31	32.00
SD	2.03	2.74	2.50	2.62	2.70	2.64	2.57	2.35	2.56	2.82	10.98

TABLE 18

Target Game Item by Item and Item by Total Correlations, Item Means and Standard Deviations (n=129).

Goldsboro, N. C.

item	1	2	3	4	5	6	7	8	9	10	total
1	1.00	.03	.11	.18	.11	-.03	.15	.09	.11	-.02	.425
2	.03	1.00	.11	-.08	.13	.05	.04	-.05	-.02	.08	.311
3	.11	.11	1.00	.13	.18	.02	.10	.06	.21	.10	.470
4	.18	-.08	.13	1.00	.18	.00	.06	.13	.07	-.10	.374
5	.11	.13	.18	.18	1.00	-.03	.00	.11	.16	.09	.477
6	-.03	.05	.02	.00	-.03	1.00	-.05	.09	.08	.15	.336
7	.15	.04	.10	.06	.00	-.05	1.00	.00	.15	.11	.377
8	.09	-.05	.06	.13	.11	.09	.00	1.00	.20	.08	.431
9	.11	-.02	.21	.07	.16	.08	.15	.20	1.00	.13	.501
10	-.02	.08	.10	-.10	.09	.15	.11	.08	.13	1.00	.390
total	.425	.311	.470	.374	.477	.336	.377	.431	.501	.390	
mean	3.44	3.97	3.44	3.67	3.52	3.72	3.71	3.64	3.53	3.39	36.03
SD	3.08	2.94	2.78	2.86	3.17	3.28	3.02	3.14	2.88	2.88	12.27

/cc:

TABLE 19

Target Game Item by Item and Item by Total Correlations, Item and Total Means and Standard Deviations for (n=260).

Owensboro, Ky.

item	1	2	3	4	5	6	7	8	9	10	total
1	1.00	.18	.13	.16	.25	.25	.12	.22	.10	.08	.576
2	.18	1.00	.21	.11	.19	.06	-.06	.14	.04	.09	.455
3	.13	.21	1.00	.07	.05	.04	.01	.08	.05	.11	.397
4	.16	.11	.07	1.00	.19	.09	.03	.14	.13	.05	.465
5	.25	.19	.05	.19	1.00	.20	.02	.06	.02	.07	.469
6	.25	.06	.04	.09	.20	1.00	.10	.05	.02	-.01	.430
7	.12	-.06	.01	.03	.02	.10	1.00	.08	.00	.00	.324
8	.22	.14	.08	.14	.06	.05	.08	1.00	.04	.17	.453
9	.10	.04	.05	.13	.02	.02	.00	.04	1.00	.04	.327
10	.08	.09	.11	.05	.07	-.01	.00	.17	.04	1.00	.378
total	.576	.455	.397	.465	.469	.430	.324	.453	.327	.378	
mean	3.08	3.48	3.19	3.43	3.03	3.19	3.31	3.12	3.18	3.08	32.09
SD	2.53	2.65	2.49	2.70	2.47	2.66	2.82	2.54	2.53	2.68	11.18

TABLE 20

Target Game Item by Item and Item by Total Correlations, Item and Total Means and Standard Deviation for (n=41).

Omaha, Neb.

item	1	2	3	4	5	6	7	8	9	10	total
1	1.00	.04	-.05	.25	-.02	.12	.03	-.02	.18	.15	.341
2	.04	1.00	.23	.10	.07	.14	.26	.29	.00	.04	.515
3	-.05	.23	1.00	.15	-.14	-.22	.29	.30	-.02	-.01	.351
4	.25	.10	.15	1.00	.19	-.06	.07	.18	.14	.19	.532
5	-.02	.07	-.14	.19	1.00	.10	.18	.12	-.09	.22	.381
6	.12	.14	-.22	-.06	.10	1.00	.15	-.04	.11	.12	.342
7	.03	.26	.29	.07	.18	.15	1.00	.25	-.07	.10	.533
8	-.02	.29	.30	.18	.12	-.04	.25	1.00	.00	.02	.506
9	.18	.00	-.02	.14	-.09	.11	-.07	.00	1.00	.20	.325
10	.15	.04	-.01	.19	.22	.12	.10	.02	.20	1.00	.466
total	.341	.515	.351	.532	.381	.342	.533	.506	.325	.466	
mean	3.00	3.32	4.00	4.07	3.22	3.76	3.54	3.41	3.41	3.17	34.90
SD	1.88	2.60	2.33	2.95	2.39	2.75	2.56	2.66	2.41	2.51	10.89

APPENDIX B

Extensive internal analysis of the data already available is a prerequisite to determining what aspects of the Target Game format can be sacrified without losing import and information. Since extensive internal analysis is necessary prerequisite to:

1. substantively meaningful external analyses
2. enlightened recommendation with respect to how the target game can and should be used.
3. informed evaluation of the utility of the current physical layout, and
4. a rational basis for debiasing the global discrepancy score and disambiguating its components,

the statistician's suggestions are discussed below.

In general, one wants to know:

Step 1. Keypunch and verify master deck. The master deck is to contain all the information available on the data sheets obtained in the Summer 1972 target game studies (550+ subjects). Each card is to include:

a) The eighteen columns of identification which currently is coded on the data sheet (incorporates school, grade, sex, race, etc.).

b) any verbal behavior? one column (1 = yes/2 = no)

c) 44 columns consisting of 22 two column fields representing the eleven pairs of calls ('chooses') and results ('hits'). (code, etc.)

d) study and deck ID

Step 2. Using the high-speed computer, compute and punch working deck. The input is the master deck. Error check and recode calls and results according to the following schedule:

old	new
0 preceded by 1-5	0
0 preceded by 23, 6-9	6
1 - 5	1 - 5
23	7
6 - 9	8 - 11
25	13
10 - 13	14 - 17
Other	ERROR

The result of this recode is a working deck parallel in format to the master deck, but with calls and results coded acording to the target region code illustrated schematically by Figure B. Other orderings will prove more useful at a later stage, but these will do for the moment.

Step 3. Compute print, and punch basic initial summary tables. On each card of the working deck there are, after the first result, 10 successive pairs of results and subsequent calls. In the 550+ cards there are a total of 5500+ such pairs of two types: a) success (the result is in the same bax as the preceding call and is equal in value for codes 2, 8, and 14, or greater or equal in value for all other codes), b) nonsuccess (all other situations). What is wanted are two cross-tabulations, results by subsequent calls (18 x 18),

one table for the situation in which results were successes in relation to the previous call, and the for the situation in which they were not. Once these basic tables are accumulated, adding them together yields a total table which we also want. These three 'full target' tables are each to be collapsed in two ways:

a) 'between box', (3 x 3) calls and results both collapsed according to: (0-5 = 0/6-11 = 1/ 12-17 = 2).

b) 'within box', (6 x 6), calls and results both collapsed according to: (0, 6, 12 = 0/ 1, 7, 13 = 1/2, 8, 14 = 2/ ..../ .5, 11, 17 = 5).

Moreover, for each of these nine tables, for each result value, the number of occurances of such a result value is the denominator with respect to which percentages of call values, as well as accumulated percentages of calls of less or equal value are to be computed. These nine tables, with frequencies, percentages, and accumulated percentages, constitute the output of Step 3 (For the three 'full target' tables, one may want in addition the percentages accumulated according to an order such as that displayed in Figure C, or some other ordering more nearly like the values actually on the target.)

The nine tables obtained from Step 3 will display the components of the target game instrument, i.e. the minimum summary information a psychometrician will need to examine (and he may wish to examine other information also) as a basis for recommendation as to how discrepancy, etc. should be scored.

Step 4. Contemplate the data for a couple of weeks. The psychometrician will be looking at the basis initial tables from the point of view of making recommendations about scoring over-estimation, discrepancy, and up-after-success-down-after-failure in each of the full target, between box and within box situations. In particular, the scoring procedure recommended is supposed to make values arising under different contingencies comparable. While this is happen-

ing, the principal researcher may wish to contemplate breakdowns of some of the tables by sex, by grade, and by race (omit District 11); if so, let the person doing Step 3 know this far in advance. The output of this contemplation is to be knowing the target game behaviors whereby the sexes, races, and grades are different from each other. e.g. does the manner in which girls and boys are different the same as the manner in which second graders and third graders are different?

Step 5. What happens here depends a lot on what the data says and little on the analyst's style. On would lend to considering each previous call ('chooses') separately, the cumulative distribution and percentile of the subsequent trials. The ordering in Figure C may be more appropriate than that of Figure B for the 'full target' situation. These percentiles constitute scoring values of calls that are comparable for different values of calls that are comparable for different values of previous results ('hits'). Folding at the median yields an analogue of discrepancy scoring values made comparable for different previous results.

Similarly, reversed percentiles after success, but not failure, gives something analogous to the up-after-success-down-after-failure score. These are considerations, not recommendations: looking at the data is the recommendation.

Whatever the analyst's style, the output of Step 5 may well be scoring keys for 'full', 'between' and 'within' overestimates, discrepancy, and up/down scales. Each such key is a table giving the recommended value to assign a call given the situation in which it occurs.

Step 6. Machine score, producing summary decks, using the scoring procedures of step 5, and the working deck, produce summary deck consisting of 18 columns ID, the developed score totaled over the ten pairs of 'hits' and 'chooses', also

FIGURE B
THE TARGET

Correspondence between working deck
and target region

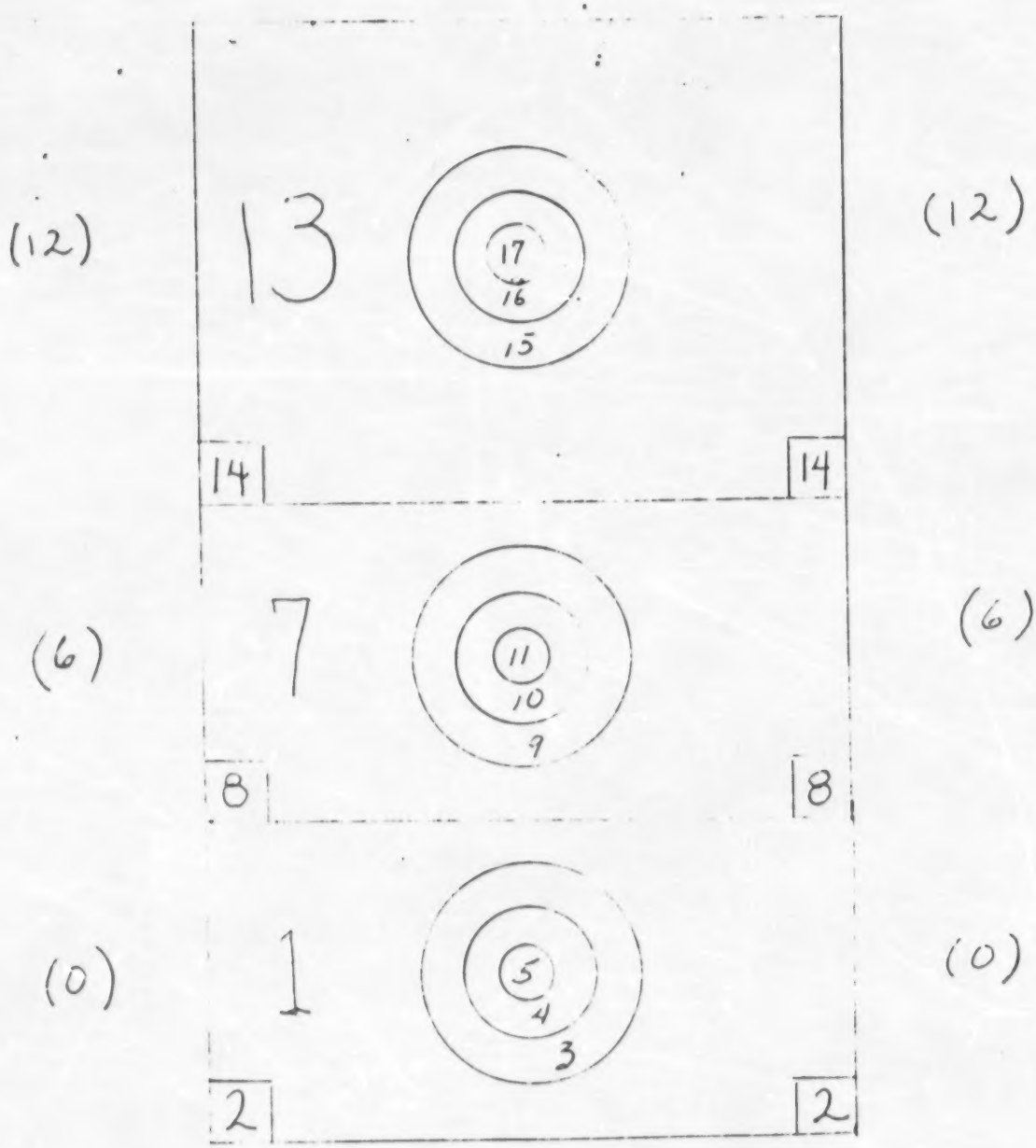
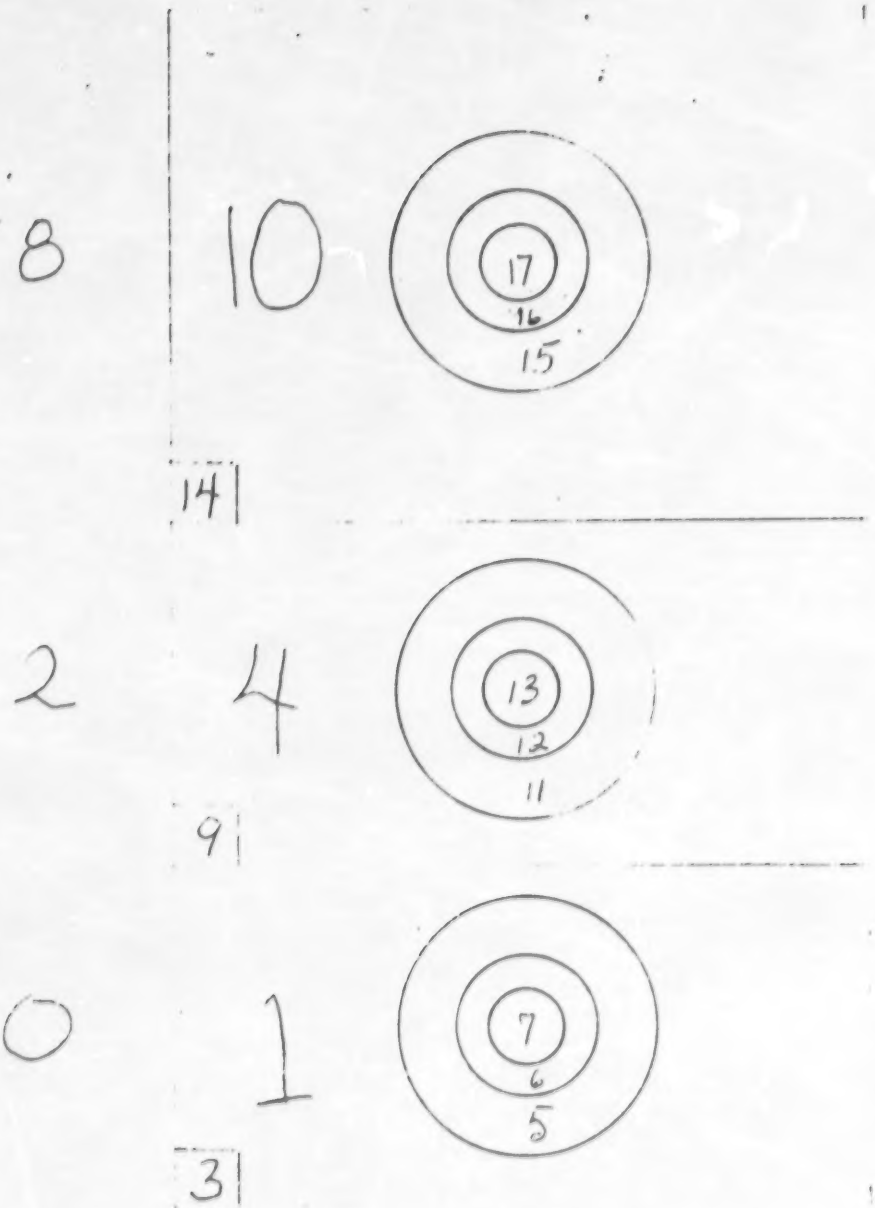


FIGURE C



scoring other target game behaviors total reward number of trials rewarded, talk, 2-8-14 count (=original 2-6-10 count) and other good ideas available at the time.

Step 7. The relationship among variables intercorrelations among variables, and whatever related analyses are helpful. Since sex and grade are in the ID field, it's possible to sneak these two variables in as zero/one without much trouble. Trying to sneak in race and age too would cause trouble in the sense of missing data. The purpose of this phase is to come to grips with the relation of six special indices concerned with call relative to previous result to the other variables in the target game. Also one wants frequence distributions for these variables. These steps would constitute the internal analysis project.

Although the plan outlined above may doubtless be enhanced, it is reasonable in its essentials as it stands. The key peg of the plan is in the pooling of items in Step 3. The 5500+ pairs so pooled run into 18 and 36 ways splits in steps 3-5 for an average of 300+ pairs for each of the 18 or 36 situations, sometimes less. Splitting the data according to any additional criterion would leave us woefully short with respect to stability of analysis in Step 5.